

Copyright (c) 1993 - 2006 Biocceleration Ltd.	GenCore version 5.1.9	94.4	2.7	969	10	US-10-363-483A-2-26081	Sequence 26081, A
OM nucleic - nucleic search, using sw model		94.4	2.7	969	10	US-10-363-483A-2-26082	Sequence 26082, A
Run on:	June 15, 2006, 03:46:50 : Search time 3642 Seconds (without alignments) 11832.892 Million cell updates/sec	94.6	2.6	8056	9	US-10-363-483A-2-26082	Sequence 26082, A
Title:	US-10-018-105A-3	94.6	2.6	8056	9	US-10-363-483A-2-26086	Sequence 26086, APP
Perfect score:	3523	90	2.6	3996	6	US-10-363-486-442	Sequence 442, APP
Sequence:	1 tatataattatataatata.....tcaattttatcaaataatgc 3525	86.6	2.5	8056	9	US-10-473-126-240	Sequence 240, APP
Scoring table:	IDENTITY_NUC_Gap0 10.0 , Gapext 1.0	84.8	2.4	8056	9	US-10-473-126-240	Sequence 240, APP
Total number of hits satisfying chosen Parameters:	37784340	82.2	2.3	7676	7	US-10-240-485-151	Sequence 151, APP
Minimum DB seq length: 0		82.2	2.3	1134	11	US-10-932-12A-166160	Sequence 166160, APP
Maximum DB seq length: 2000000000		81.2	2.3	3089	10	US-10-954-778-150	Sequence 150, APP
Post-processing: Minimum Match 0%		81.2	2.3	4985	6	US-10-094-240-10	Sequence 10, APP
Post-processing: Maximum Match 100%		81.2	2.3	4985	7	US-10-056-405-10	Sequence 10, APP
Post-processing: Listing first 45 summaries		81.2	2.3	3931	6	US-10-006-780-1	Sequence 1, APP
Database :		35	2.3	8759	9	US-10-473-126-232	Sequence 232, APP
Published Applications NA_Main:*		35	2.3	8759	9	US-10-473-126-250	Sequence 250, APP
1: /EMC_Celerra_SIDS3/ptodata/2/pubnra/US07_PUBCOMB.seq;*		36	2.2	8759	9	US-10-473-126-378	Sequence 378, APP
2: /EMC_Celerra_SIDS3/ptodata/2/pubnra/US09A_PUBCOMB.seq;*		37	2.2	8759	9	US-10-363-34A-8883	Sequence 8883, A
3: /EMC_Celerra_SIDS3/ptodata/2/pubnra/US09B_PUBCOMB.seq;*		38	2.2	8759	9	US-10-363-34A-8884	Sequence 8884, A
4: /EMC_Celerra_SIDS3/ptodata/2/pubnra/US09C_PUBCOMB.seq;*		39	2.2	8759	9	US-10-363-43A-28883	Sequence 28883, A
5: /EMC_Celerra_SIDS3/ptodata/2/pubnra/US09C_PUBCOMB.seq;*		40	2.2	535	9	US-10-311-45F-1221	Sequence 1221, AP
6: /EMC_Celerra_SIDS3/ptodata/2/pubnra/US10A_PUBCOMB.seq;*		41	2.2	535	10	US-10-363-43A-28884	Sequence 28884, A
7: /EMC_Celerra_SIDS3/ptodata/2/pubnra/US10B_PUBCOMB.seq;*		42	2.2	535	10	US-10-363-43A-28884	Sequence 28884, A
8: /EMC_Celerra_SIDS3/ptodata/2/pubnra/US10C_PUBCOMB.seq;*		43	2.2	535	10	US-10-311-45F-1221	Sequence 1221, AP
9: /EMC_Celerra_SIDS3/ptodata/2/pubnra/US10D_PUBCOMB.seq;*		44	2.2	5163	7	US-10-311-45F-1221	Sequence 1221, AP
10: /EMC_Celerra_SIDS3/ptodata/2/pubnra/US10E_PUBCOMB.seq;*		45	2.2	564	8	US-10-437-963-7175	Sequence 7175, AP
11: /EMC_Celerra_SIDS3/ptodata/2/pubnra/US10F_PUBCOMB.seq;*							
12: /EMC_Celerra_SIDS3/ptodata/2/pubnra/US10G_PUBCOMB.seq;*							
13: /EMC_Celerra_SIDS3/ptodata/2/pubnra/US11A_PUBCOMB.seq;*							
14: /EMC_Celerra_SIDS3/ptodata/2/pubnra/US11B_PUBCOMB.seq;*							
15: /EMC_Celerra_SIDS3/ptodata/2/pubnra/US11C_PUBCOMB.seq;*							
16: /EMC_Celerra_SIDS3/ptodata/2/pubnra/US11D_PUBCOMB.seq;*							
Post. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.							
SUMMARIES							
Result No.	Score	Query	Match Length	DB ID	Description	Qy	338 ATTCGCCAGTGGAGTAATCAATTAACTACTATAACCCATTCTCTAGCT 397
							1 ATTCGCCAGTGGAGTAATCAATTAACTACTATAACCCATTCTCTAGCT 60
1	2769	78.6	2769	7	US-10-032-585-6675	Qy	398 GATGAATTTGGATGATCATGAAATCAATTAACTATACTATAACCCATTCTCTAGCT 457
2	408.8	11.6	2571	11	US-10-932-182A-5831	Qy	61 GATGAATTTGGATGATCATGAAATCAATTAACTATACTATAACCCATTCTCTAGCT 120
3	403.8	11.5	2574	11	US-10-932-182A-2281	Db	458 GATGAATTTGGATGATCATGAAATCAATTAACTATACTATAACCCATTCTCTAGCT 517
4	396	11.2	2580	11	US-10-932-182A-78462	Qy	121 GATGAATTTGGATGATCATGAAATCAATTAACTATACTATAACCCATTCTCTAGCT 180
5	383	10.9	2577	11	US-10-932-182A-6424	Qy	518 CAAACACATCACAGATGATCATGAAATCAATTAACTATACTATAACCCATTCTCTAGCT 577
6	179.6	5.1	1992	6	US-10-128-714-7563	Db	181 CAAACACATCACAGATGATCATGAAATCAATTAACTATACTATAACCCATTCTCTAGCT 240
7	179.6	5.1	1992	6	US-10-128-714-7563	Qy	578 TCCTCAATTAGAAATCAATTAACTATACTATAACCCATTCTCTAGCT 637
8	127.7	3.6	2910	11	US-10-932-182A-77545	Db	241 TCCTCAATTAGAAATCAATTAACTATACTATAACCCATTCTCTAGCT 300
9	117.6	3.3	462	3	US-09-864-40RA-2265	Qy	638 TATCGGGCGGCGATGATCATGAAATCAATTAACTATACTATAACCCATTCTCTAGCT 697
10	117.6	3.3	8056	9	US-10-473-126-316	Qy	Sequence 10, APP
11	116.4	3.3	688	9	US-10-653-047-703	Db	301 TATCGGGCGGCGATGATCATGAAATCAATTAACTATACTATAACCCATTCTCTAGCT 360
12	116.2	3.3	2910	11	US-10-932-182A-7077	Qy	698 TCATTAGTAAAGTTTAAAGATATACTAGACAACTTCTATCAT 757
13	113.6	3.2	2048	6	US-10-128-714-6563	Qy	Sequence 2048, APP
14	113.6	3.2	2048	6	US-10-128-714-6563	Qy	Sequence 447, APP
15	113.6	3.2	4048	6	US-10-128-714-563	Qy	Sequence 1725, APP
16	113.6	3.2	4048	6	US-10-128-714-563	Qy	Sequence 2083, APP
C 17	106.8	3.0	3089	10	US-10-954-778-160	Qy	Sequence 2082, A
18	106.8	3.0	4985	6	US-10-047-703	Db	Sequence 2082, A
C 19	106.8	3.0	4985	7	US-10-056-405-10	Qy	
20	106.2	3.0	575	9	US-10-653-047-447	Qy	
21	97.4	2.8	286	6	US-10-425-11-17269	Qy	
C 22	94.4	2.7	963	9	US-10-363-345A-26081	Qy	
23	94.4	2.7	969	9	US-10-363-345A-26082	Qy	

ALIGNMENTS

RESULT 1
US-10-012-585-6675
; Sequence 6675, Application US/10032585
; Publication No. US2003180933A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10181-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 6675
; LENGTH: 2769
; ORGANISM: Candida albicans
US-10-012-585-6675

Query Match 78.6%; Score 2769; DB 7; Length 2769;
Best Local Similariy 100.0%; Pred. No. 0;
Matches 2769; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 338 ATTCGCCAGTGGAGTAATCAATTAACTACTATAACCCATTCTCTAGCT 397

Db 1 ATTCGCCAGTGGAGTAATCAATTAACTACTATAACCCATTCTCTAGCT 60

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2006, 14:08:32 : Search time 52 Seconds

(without alignments)
1551.985 Million cell updates/sec

Title: US-10-018-105A-4

Perfect score: 420

Sequence: 1 MSDESYTQNNTNQPIRPS.....LGKXHGGNKSISFPNPKY 922

Scoring table: BLOSUM62

Gapop 10.0 . Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 0%

Listing first 45 summaries

Database : Issued Patents_AA:*

1: /EMC_Celerra_STDS3_ptodata/2/ias/5_COMB.pep:*

2: /EMC_Celerra_STDS3_ptodata/6/ias/6_COMB.pep:*

3: /EMC_Celerra_STDS3_ptodata/2/ias/7_COMB.pep:*

4: /EMC_Celerra_STDS3_ptodata/2/ias/R_COMB.pep:*

5: /EMC_Celerra_STDS3_ptodata/2/ias/PCTUS_COMB.pep:*

6: /EMC_Celerra_STDS3_ptodata/2/ias/RE_COMB.pep:*

7: /EMC_Celerra_STDS3_ptodata/2/ias/backfilesl1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	3785	78.5	785	2 US-09-248-796A-2431	Sequence 20431, A
2	1680	34.9	659	2 US-09-538-092-1717	Sequence 717, App
3	834	17.3	814	2 US-09-248-796A-20405	Sequence 20405, A
4	716	14.9	168	2 US-09-248-796A-20430	Sequence 20430, A
5	251	5.2	326	2 US-09-502-540-11640	Sequence 11640, A
6	172.5	3.6	1091	2 US-09-538-092-169	Sequence 469, App
7	166.5	3.5	2983	2 US-10-177-502-4	Sequence 4, App1
8	166	3.4	719	2 US-09-48-796A-7559	Sequence 193, App
9	161	3.3	366	2 US-09-712-363-193	Sequence 168, App
10	161.5	3.3	471	2 US-09-134-001C-5106	Sequence 5106, App
11	159.5	3.3	1177	2 US-09-134-001C-5106	Sequence 18, App1
12	159.5	3.3	1253	2 US-09-958-617A-18	Sequence 18, App1
13	159.5	3.3	1584	2 US-09-457-040B-27	Sequence 17, App1
14	159.5	3.3	2150	3 US-10-115-322-17	Sequence 15, App1
15	158.5	3.3	785	2 US-09-88-538B-348	Sequence 24, App1
16	158	3.3	651	1 US-08-431-050-24	Sequence 24, App1
17	158	3.3	651	1 US-08-938-534-24	Sequence 24, App1
18	158	3.3	651	2 US-09-345-294-24	Sequence 24, App1
19	157	3.3	858	2 US-09-248-796A-19055	Sequence 19055, A
20	154	3.2	1259	2 US-09-949-016-10365	Sequence 10365, A
21	151.5	3.2	982	2 US-09-248-796A-0628	Sequence 20628, A
22	152	3.2	2870	2 US-09-79-467A-15	Sequence 15, App1
23	152	3.2	2870	2 US-09-655-160-15	Sequence 15, App1
24	152	3.2	3178	2 US-09-479-467A-4	Sequence 4, App1
25	152	3.2	3178	2 US-09-655-160-4	Sequence 4, App1
26	151	3.1	315	2 US-09-710-279-2100	Sequence 2100, Ad
27	151	3.1	324	2 US-09-134-001C-080	Sequence 498, Ad
28	151	3.1	10182	2 US-09-134-001C-119	Sequence 315, Ad
29	150	3.1	676	2 US-09-134-001C-118	Sequence 431, Ad
30	147	3.0	404	2 US-09-710-279-398	Sequence 398, Ad
31	147	3.0	5024	2 US-09-710-279-2964	Sequence 2962, Ad
32	146.5	3.0	1444	2 US-09-949-016-9652	Sequence 9652, Ad

ALIGNMENTS

RESULT 1	US-09-248-796A-20431	Score: 3785; DB 2; Length 785;
1	Sequence 20431, Application US/0948796A	
2	Patent No. 6747137	
3	GENERAL INFORMATION:	
4	APPLICANT: Keith Weinstock et al	
5	TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS	
6	FILE REFERENCE: 107196_132	
7	CURRENT APPLICATION NUMBER: US/09/248-796A	
8	CURRENT FILING DATE: 1999-02-12	
9	PRIOR APPLICATION NUMBER: US 60/074,725	
10	PRIOR FILING DATE: 1998-02-13	
11	PRIOR APPLICATION NUMBER: US 60/096,409	
12	PRIOR FILING DATE: 1998-08-13	
13	NUMBER OF SEQ ID NOS: 28208	
14	SEQ ID NO: 28208	
15	LENGTH: 785	
16	TYPE: PRT	
17	ORGANISM: Candida albicans	
18	US/09-248-796A-20431	
19	Query Match 78.5%; Best Local Similarity 96.2%;保守 Matches 740; Conservative Matches 740; Mismatches 2; Mismatches 13; Indels 14; Gaps 7;	
20	Qy 1 MDPDSQTGNSSTGPIRSDEVDQDHNQITNDQCAISDSDDELELSASEVSEVKSE- Db 18 MSDBSYTSNTNSTNPILSDEVDLDRNQITNDQCAISDSDDELEKPELESEVSEVKQ 77	
21	Qy 60 ---QQQHQETTSNDNAKPLTKSGSSTSKKSNLTDKORTINPMSLGGDDTINSKHRRN 116 Db 78 QQQQHQETTSNDNAKPLTKSGSSTSKKSNLTDKORTINPMSLGGDDTINSKHRRN 137	
22	Qy 117 YMASSLKDPLKQTDNNTNTHAIPITPPIPITNNANKRSQSLNLPPLIK 176 Db 138 YMSSRLKDPLKQTDNNTNTHAIPITPPIPITNNANKRSQSLNLPPLIK 197	
23	Qy 177 KTTIGRNNSNNPENDLYSPMTKSTDNTDNT-TSTTAHHMKGATGTLGVGCTTA 235 Db 198 KTTIGRNNSNNPENDLYSPMTKSTDNTDNT-TSTTAHHMKGATGTLGVGCTTA 257	
24	Qy 234 ATATAAA ---GRPRSSIDSDEADSHASRSQTEEDCPATPGDHRNGIDFEDDFP 292 Db 258 ATATATAGGGRRPSRISDEADSHASRSQTEEDACPATPGDHRNGIDFEDDFP 317	
25	Qy 293 IRRERAYLQOKONTAKTILRIRERONLSKNTTSCASRPVYHMKNNKKANGEDS-GG 351 Db 318 IRRERAYLQOKONTAKTILRIRERONLSKNTTSCASRPVYHMKNNKKANGEDS-GG 377	
26	Qy 332 SSMALKYTPRNLIKURKTSRFEPTHEINSSSEYLTKQOPPYDQQLSLTSSTST 411 Db 378 SSMALKYTPRNLIKURKTSRFEPTHEINSSSEYLTKQOPPYDQQLSLTSSTST 437	
27	Qy 412 --SGGSQSQWFGARSDGNSLDDPSLPHSESETIAPDPSLVPSQSRDFP 469 Db 438 SGGSQSQWFGARSDGNSLDDPSLPHSESETIAPDPSLVPSQSRDFP 497	
28	Qy 470 RNEGETWMLDCPTDSEMCLAKAFCHPLAEIRDQETREVKLFPSYFCPHTE 529	

GenCore version 5.1.9
(c) 1993 - 2006 Biacceleration Ltd.

CM protein - protein search, using sw model

Run on: June 13, 2006, 14:20:38 : Search time 182 Seconds
(without alignments)
2346,616 Million cell updates/sec

Title: US-10-018-105A-4
Perfect score: 4820
Sequence: 1 MSDSESYQNNTNQPIRS.....LGKKGNGNTKTSISPPNKE 922

Scoring table: BLOSUM62
Gapext 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database : Published_Applications_JA_Main.*

1: /EMC_Celerra_SIDS3/podata/2/pubpaa/US07_PUBCOMB.pep.*

2: /EMC_Celerra_SIDS3/podata/2/pubpaa/US08_PUBCOMB.pep.*

3: /EMC_Celerra_SIDS3/podata/2/pubpaa/US10_PUBCOMB.pep.*

4: /EMC_Celerra_SIDS3/podata/2/pubpaa/US10B_PUBCOMB.pep.*

5: /EMC_Celerra_SIDS3/podata/2/pubpaa/US11_PUBCOMB.pep.*

6: /EMC_Celerra_SIDS3/podata/2/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	4814	922 4	US-10-012-585-7675	7675	AP	Sequence 7675, AP
2	1679	34.8	859 2	US-0-945-749-1	App1	Sequence 2, App1
3	1666	34.6	858 2	US-0-945-749-2	App1	Sequence 2, App1
4	1089	22.8	663 4	US-10-128-714-3	563	Sequence 3563, AP
5	1089	22.8	663 4	US-10-128-714-3	563	Sequence 3563, AP
6	856	17.8	969 2	US-0-945-749-3	App1	Sequence 3, App1
7	3455	7.2	154 4	US-0-945-884-8	2066	Sequence 2066, AP
8	298	6.2	94 2	US-10-45-115-321932	321932	Sequence 321932, AP
9	298	6.2	387 2	US-0-945-749-8	App1	Sequence 8, App1
10	234.5	4.9	1198 6	US-11-097-143-7	905	Sequence 7905, AP
11	213	4.4	380 2	US-0-945-749-3	App1	Sequence 3, App1
12	213.5	4.4	386 4	US-10-156-761-9	574	Sequence 9574, AP
13	182.5	4.2	371 4	US-10-156-761-10	664	Sequence 10664, AP
14	182.5	3.8	373 3	US-10-156-761-10	664	Sequence 10664, AP
15	179.5	3.7	373 3	US-11-097-143-248	App1	Sequence 48, App1
16	179.5	3.7	1172 4	US-10-451-674-574	App1	Sequence 574, App1
17	179.5	3.7	2344 5	US-10-156-723-13	3978	Sequence 13978, A
18	179	3.7	354 6	US-10-156-761-13	3791	Sequence 13791, A
19	179	3.7	1199 6	US-11-097-143-16	64	Sequence 1664, AP
20	175.5	3.6	1246 6	US-11-097-143-16	64	Sequence 1164, AP
21	172.5	3.6	328 5	US-10-72-923-9	311	Sequence 931, AP
22	171.2	3.6	2165 5	US-10-72-923-13	347	Sequence 1347, A
23	170.5	3.5	343 4	US-10-156-761-14	171	Sequence 14171, A
24	168.9	3.5	1468 6	US-11-097-143-22	779	Sequence 22779, A
25	168.5	3.5	816 5	US-10-457-555-32	App1	Sequence 32, App1
26	166.5	3.5	1046 6	US-11-165-819-6	App1	Sequence 6, App1
27	166.5	3.5	1046 6	US-11-165-819-6	App1	Sequence 26, App1
28	166.5	3.5	2261 5	US-10-410-048-60	App1	Sequence 60, App1
29	166.5	3.5	2283 4	US-10-127-502-4	App1	Sequence 4, App1
30	166.5	3.5	2283 6	US-11-032-585-7876	App1	Sequence 7876, AP
31	166	3.4	2112 6	US-11-097-143-8001	App1	Sequence 8001, AP
32	165	3.4	4498 4	US-10-712-124-68	App1	Sequence 68, App1

RESULT 1
US-10-032-585-7675

: Sequence 7675, Application US/10032585

: Publication No. US2003018095A1

: GENERAL INFORMATION:

; APPLICANT: Perry, Ronner D.

; APPLICANT: Bo, Jiang

; APPLICANT: Howard, Bussey

; FILE REFERENCE: 10182-005-939

; CURRENT APPLICATION NUMBER: US/10/032,585

; NUMBER OF SEQ ID NOS: 0000

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO: 7675

; LENGTH: 922

; TYPE: PRT

; ORGANISM: Candida albicans

; US-10-032-585-7675

Query Match 99.9%; Score 4814; DB 4; Length 922;

Best Local Similarity 99.9%; Pred. No. 0; Mismatches 0; Gaps 0;

Matches 921; Conservative 0; APPLICANT: Perry, Ronner D.

APPLICANT: Bo, Jiang

APPLICANT: Howard, Bussey

FILE REFERENCE: 10182-005-939

CURRENT APPLICATION NUMBER: US/10/032,585

NUMBER OF SEQ ID NOS: 0000

SOFTWARE: PatentIn version 3.1

SEQ ID NO: 7675

LENGTH: 922

TYPE: PRT

ORGANISM: Candida albicans

US-10-032-585-7675

Query Match 99.9%; Score 4814; DB 4; Length 922;

Best Local Similarity 99.9%; Pred. No. 0; Mismatches 0; Gaps 0;

Matches 921; Conservative 0; APPLICANT: Perry, Ronner D.

APPLICANT: Bo, Jiang

APPLICANT: Howard, Bussey

FILE REFERENCE: 10182-005-939

CURRENT APPLICATION NUMBER: US/10/032,585

NUMBER OF SEQ ID NOS: 0000

SOFTWARE: PatentIn version 3.1

SEQ ID NO: 7675

LENGTH: 922

TYPE: PRT

ORGANISM: Candida albicans

US-10-032-585-7675

Query Match 99.9%; Score 4814; DB 4; Length 922;

Best Local Similarity 99.9%; Pred. No. 0; Mismatches 0; Gaps 0;

Matches 921; Conservative 0; APPLICANT: Perry, Ronner D.

APPLICANT: Bo, Jiang

APPLICANT: Howard, Bussey

FILE REFERENCE: 10182-005-939

CURRENT APPLICATION NUMBER: US/10/032,585

NUMBER OF SEQ ID NOS: 0000

SOFTWARE: PatentIn version 3.1

SEQ ID NO: 7675

LENGTH: 922

TYPE: PRT

ORGANISM: Candida albicans

US-10-032-585-7675

ALIGNMENTS

; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus
 ; TITLE OF INVENTION: Methods of Use
 ; FILE REFERENCE: 10182-018-999
 ; CURRENT APPLICATION NUMBER: US 10/128,714
 ; CURRENT FILING DATE: 2002-04-23
 ; PRIORITY NUMBER: US 60/285,697
 ; PRIORITY FILING DATE: 2001-04-21
 ; PRIORITY NUMBER: US 60/287,066
 ; PRIORITY FILING DATE: 2001-04-27
 ; PRIORITY NUMBER: US 60/295,890
 ; PRIORITY FILING DATE: 2001-06-05
 ; PRIORITY NUMBER: US 60/303,899
 ; PRIORITY FILING DATE: 2001-07-09
 ; PRIORITY NUMBER: US 60/316,362
 ; PRIORITY FILING DATE: 2001-08-31
 ; NUMBER OF SEQ ID NOS: 8603
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO: 3563
 ; LENGTH: 663
 ; TYPE: PRT
 ; ORGANISM: Aspergillus fumigatus
 ; US-10-128-714-3563

Query Match 22.8% Score 1099; DB 4; Length 663;
 Best Local Similarity 32.7%; Pred. No. 3..-67;
 Matches 292; Conservative 98; Mismatches 241; Indels 262; Gaps 22;

Qy 4 SESYQONSTNQTPRSPDEVLDRHNRNITDCAISDSEBLEKSLLESEVVKSKERQQOH 63
 Db 7 SEKYMCLTRPSTPVE---LDDDRFLP-----DSPPRIATADI-SLSQNTAQAH 54

Qy 64 HOETTSNAKPLTRKGSGSSTIKSNLTDKDRINTPMSLSCGDDTINSGHKNRNNMNSR 123
 Db 55 HOETPQ-----RDLISLSTODAREAGSLSDPQAIIID 87

Qy 124 KDPYIJKD-----NTDNDSTNH----THLAIPPIPPIPITANNSRK--SQ 167
 Db 88 DRSQKDINTGARRPSVDPRNGVNRHTGWTSHQL-----ANNRESFSA 134

Qy 168 LENLPPLKKTKTIGRNNSNFFNDLSPMTKMTNDSEDTITSPTANMKLGATTTG 227
 Db 135 RSSSP-----NSVEAFADPREFREANTLE-----SHAAPOLEILQRTVSG 176

Qy 228 VGTGT-TATATAAAGRGRSSSIDEADHSRSQTEEDVCPMGDHRVNGIDP 286
 Db 177 ---GTHPRPFSNAIRPQGDQLEPN-----DESCVPTEQPIRSPVIDY 222

Qy 287 DEIDEPFIREBAYLQKQITAKNLRIDFQNLKNTGASHRUP---YHHHSNNRKK 343
 Db 223 EELBEFVA-----LSRQPKPSTSRR-----KOSUSSSRSPRVYDRLRQLRS 266

Qy 344 NGGGGGGSSMAALKTPKNLKLTLSRPEPTENSSSEPVYKLTQOPPKYDQQLS 403
 Db 267 DVBGKEKRSAA-----DRSSPDMDADLKTDKYYANVDRKD 304

Qy 404 LTSETSETSGSGSQCQVKFGARISDINGSLPDRSLPHSESETHTADPILSPVSPECQ 463
 Db 305 -----TVEKLQNENETPTRGFSSSESQSVHAEGLDVLFLP 342

Qy 464 SVQDFRNGEE-TWIDCTCPDSENMLAKAGTHPTADIRQETREVFLPSY 521
 Db 343 TFRDLFQGZPGGGWWLDVNPTEBVAALSRSRSPHPUTEDLTQEARKEVLPQY 402

Qy 522 FVCPIHPF-EAKKESDLYLEPTINYIVFHGILTFSPDSHAPAYREVROLRDYDVS 580
 Db 403 FUCPCTPYQDQKTERPMEPMVYVUPDCUSSFETNPBANTRIGLRVSYIS 462

Qy 561 ADWCVYLIDITDITGAPVINGEYEAADATEDVFTARDYDPSMLOIGSERRYVTLN 640
 Db 463 SDWICYAMDDIVDSFSPVREIESEALDVFTARYDDESPFLPRIGGRKXMSLM 522

GENERAL INFORMATION:
 ; APPLICANT: Liang, Bo
 ; APPLICANT: Hu, Wengi
 ; APPLICANT: Tishkoff, Daniel
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Broshkin, Alexey M
 ; APPLICANT: Lemieux, Sébastien M

RESULT 4

US-10-128-714-3563
 Sequence 3463, Application US/10128714
 Publication No. US200301101341
 ; APPLICANT: Liang, Bo
 ; APPLICANT: Tishkoff, Daniel
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Broshkin, Alexey M
 ; APPLICANT: Lemieux, Sébastien M

Qy 541 RLLSKRADVTKMFAKRCQEERAASSGGYTORQVNLCQQOAAAPPBPNPITSPINSTNL 700
 Db 543 RLLSKRADVTKMFAKRCQEERAASSGGYTORQVNLCQQOAAAPPBPNPITSPINSTNL 700
 Qy 701 NSLGKTSRGGYGVGIGNGPMPGNTNTNTGSPSP0000QHGITNKSPPIDARP 760
 Db 542 -----QYSVT-----P 547

Qy 761 RADIALYLGDIOPHITHMNLQNLAYEKIPFSRSHSYNLAQLOVESNSNNKITEMSKITL 820
 Db 548 RDIGIYLGLDIDQHVTWMSNLAFERKMLSLSHNTYLAQLNNTNLVIGHNVNARVKLVTL 607
 Qy 821 IGMALYPLAULVGLFGMANVRYVEGEGTNLGRHFFIVGVLFLITIGSPTAQMN 873
 Db 608 IATMLYPLAULCGLFGMANVRYFBGQBCLGNFFGIVGVIAVILSGIARY 660

RESULT 5
 US-10-128-714-8563
 Sequence 8563, Application US/10128714
 Publication No. US20030119013A1
 GENERAL INFORMATION:
 TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus*
 and
 APPLICANT: Hu, Wensi
 APPLICANT: Fishhoff, Daniel
 APPLICANT: Zamudio, Carlos
 APPLICANT: Eroshkin, Alexey M
 APPLICANT: Lemieux, Sébastien M
 TITLE OF INVENTION: Methods of Use
 FILE REFERENCE: 10182-018-999
 CURRENT APPLICATION NUMBER: US/10/128-714
 CURRENT FILING DATE: 2002-04-23
 PRIOR APPLICATION NUMBER: US 60/285,697
 PRIOR FILING DATE: 2001-04-23
 PRIOR APPLICATION NUMBER: US 60/287,066
 PRIOR FILING DATE: 2001-04-27
 PRIOR APPLICATION NUMBER: US 60/295,890
 PRIOR FILING DATE: 2001-06-05
 PRIOR APPLICATION NUMBER: US 60/303,899
 PRIOR FILING DATE: 2001-07-09
 PRIOR APPLICATION NUMBER: US 60/316,362
 PRIOR FILING DATE: 2001-08-31
 NUMBER OF SEQ ID NOS: 8603
 SOFTWARE: Patentin version 3.1
 SEQ ID NO: 8563
 LENGTH: 663
 TYPE: PRT
 ORGANISM: *Aspergillus fumigatus*
 US-10-128-714-8563

Query Match 22.0%; Score: 1099; DB: 4; Length: 663;
 Best Local Similarity 32.7%; Fred. No. 3.5-67;
 Matches 292; Conservative 98; Mismatches 241; Indels 262; Gaps 22;

Qy 4 SIEYVONSTNPOTPRSDDEVLDHRNQITNDCA1SDELEKSEVYKSEKOQH 63
 Db 7 SEKYVCLTRPSPVPE---LDHFHQI---LDPDHFQI---DSPPRIETADI-SLSRONTAQHAY 54

Query Match 17.8%; Score: 856; DB: 2; Length: 969;
 Best Local Similarity 24.1%; Fred. No. 4-50;
 Matches 270; Conservative 143; Mismatches 248; Indels 460; Gaps 29;

Qy 11 STTNPOTPRSDDEVLDHRNQITNDCA1SDELEKSEVYKSEKOQH 58
 Db 44 SDSPRPTOLHDNQHNGQTDFDQI-DSWGLHESDSTDNIKSDEPSLKGAFIDR 102

Qy 59 ---KQOQHQEITSD-NAKPLRKXGSSTIKKSNLTDDRTNRS-LSGDDDNNSGH 112

Db 103 PMSQPRBGPQSVSTVQPIQMKPSTPKAGRLPSDQRNSLVSLSPE--LESWL 160

Qy 113 KRNTRNNSSLRKDEYKLKDNTDNSTNNTHLAPIPIPITPTLNAA----NKSRRK 165

Db 161 KRR---KSTVKSPVUDNSPDRQSMANNNDVVIDD-ALMHVNNSNASTGVNDSKRK 215

Qy 166 SOLENUPPLKKTYGRNANAVNENDLVSMTAKAKINDBDTNTSTANRKGIGATT 225

Db 216 ---KKR---GSDDSN-KN---SKSTSDDNDEDEYNX----- 244

Qy 226 LGVGTGTTATAAAGREPRSSRSDSEADSHAERSSSQETEDVCFPM--VGDHIVRYNG 283

Qy 305 -----IVEKLQNNEPFRGPFSSQSOTHAEGLDVLVLPED 342
 Qy 464 SVRDLFNRGEE---TMWLDCTCPTDSENKMLAKAFGHPPLTAEDIHQETREKVELFKSTY 521
 Db 343 TFRDLFOLQFPEGGYWMWDVNTPEETAAALSRAFSHTPHUETEDLTOEARKEVLFQXY 402
 Qy 522 FWCFTP-EAKESEDDYLEPNVYTYVPHGII-TPHSPSPISHANTVRVQLRDYDUS 580
 Db 403 FWCFTP-YQDOKTDORPMEPNFTAVVYFQDVLSPSTPENPHANVRKIGLRDVSLS 462

Qy 581 AHWLYALDIDTQDAPVIGIYEADEDATDVTARDTDFSMQRGESRKVATLUM 640
 Db 463 SWMCYANIDDIVDPSGFVIRETEISAEATDLYTARVDDFESLPRIGGLRKWMSLM 522
 Qy 641 RLLSGRADVIMPAKRCOEANSSSGYQROYNLQQQQQAPPNNPITSPINSTNL 700
 Db 522 RLGRADVINGFSKRCN----- 541

Qy 701 NSLGJSTSTCGVGVCINFGPMPGNNNTNTNTGSPSPFQQQQQGNTNSKFPPIPDARP 760
 Db 542 ----- 547

Qy 761 RADIALYLGDIOPHITHMNLQNLAYEKIPFSRSHSYNLAQLOVESNSNNKITEMSKITL 820
 Db 548 REDIGLYDQDHITMFPNLAYEKIPFSRSHSYNLAQLOVESNSNNKITEMSKITL 607
 Qy 821 TGTMVPLNYTGLFGAVNRYVEGEGTNLGRHFFIVGVLFLITIGSPTAQMN 873
 Db 608 LATMLVPLNLCGLFGAVNRYVEGEGTNLGRHFFIVGVLFLITIGSPTAQMN 660

RESULT 6
 US-08-945-749-3
 Sequence 3, Application US/0894549
 Publication No. US20020138880A1
 GENERAL INFORMATION:
 APPLICANT: MACLARID, Colin W
 APPLICANT: HAY, Robert J
 APPLICANT: Auckland Pastoral Agriculture Research Institut
 TITLE OF INVENTION: ALUMINUM RESISTANCE GENE
 CURRENT APPLICATION NUMBER: US/08/945-749
 CURRENT FILING DATE: 1999-01-12
 EARLIER APPLICATION NUMBER: PC/7N996/0035
 EARLIER FILING DATE: 1996-05-01
 EARLIER APPLICATION NUMBER: NZ 272039
 EARLIER FILING DATE: 1995-05-01
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: Patentin Ver. 2.0
 SEQ ID NO: 3
 LENGTH: 969
 TYPE: PRT
 ORGANISM: *Saccharomyces cerevisiae*
 US-08-945-749-3

245 -----RSSS-----LSSNNISLDVCLVLDDGSEYVKAW 275
 Db US-09-864-408A-2266
 Qy Query Match 7.2%; Score 345.5; DB 3; Length 154;
 284 IDPDEDEPRBERBAYLQLQMTAKNLRIDEQNSKUNFTSGASRHYHHSNNRK 343
 Best Local Similarity 44.6%; Pred. No. 6.3e-16;
 Matches 79; Conservative 20; Mismatches 39; Indels 39; Gaps 6;
 Db 276 PDCVLEEFPSKEETE--LRSQAI-----ODAFAFIYFOYDDE 312
 Qy 351 GSSMAALKYTPKNLKKTLSPFPEFHNSSSSEIVELTKQOPPKYDQSLTSSTSS 410
 344 NNGDGEGSSMALKYTPKNLKKTLSPFPEFHNSSSSEIVELTKQOPPKYDQSLTSSTSS 410
 Db 1 GSS-AKALYTF-----ISR---TEKNCVSHEMTVSNNES--VREDDKPPDL-----42
 313 D----GTSNEGILFSRPIVNTDVPBLGNRNVETENLNGRLPKRLAWHLIQRM 367
 Qy 411 TSGSGCOKYKGKGGARISPGING-----SLPDRFLPUSSESETHTAD 454
 404 LNTSNTSSTSGSSGCGYKGGARISPGINGSSL-----PD--RPLPHSS 446
 Db 43 ----HPDTGTRGKNGEKGENDSSYSRASYYLTQNTEQISRSFSEDETSHASD 97
 368 VL-----GSNTSKDSKRISQGLDNLYVRN1QYPHIISNPNEHFTFVRL 418
 Qy 455 IPSLVSPOQSYTDLFRNGEETWMLDCPTDSMKLAKAFGHPTADIMQEST 511
 447 EETHADPISLVSPCGSVRDF-----469
 Db 98 IFSLSBISQCPYKGDQDWWLDCSICPTDDERCIATKPGHPTADIMQEST 511
 419 DSVASPTISGLLQEPQRQDPLFASYVSDNASHKTHKPNSPPIKAETVQSQGAT 478
 Qy 470 -----RNGBET-----WMLDCCTPDSKMRKAFGTHPLTAEDIKMQETREKYE 515
 Db 479 AKNPSTLSSMANTEDVPPMWDVSNTPEBKMLSKA-----GTHPLTAEDIKMQETREKYE 538
 RESULT 8
 Db Sequence 321932, Application US/10425115
 Qy 516 LFKSITPCPHTF-----ADKES-----534
 Db 539 LPDYLTCFCSRTDVAKHYRRKERKQSBATLHESISRKSQAYGMENSNARN 598
 Qy 535 -----
 Db 599 NSTSNASRSKWLPSLARRRSANARTNTSSSSKRVKSEKKMEENEKFKRKGDRH 658
 Qy 535 -----EDYLEPINYTVFHDGLTHFPISHPANVRGRQLDYVSDWLUCLALDE 591
 Db 659 KPREGLBFLANVYTVFPTGVLTHAFTPHINVRBRALLDYUNVTSDATAVLLD 718
 Qy 592 ITDGFAPVHGKEYEADIAEDAVTAFTARDTDFSS-----624
 Db 719 ITDAPAPTELZDEVYEDIAILKMQHQSDDSSSDSDGASDEDAPPDVYKK 778
 Qy 625 -----MQRICEGSRKATWALIISGRAVIT 650
 Db 779 TSYSSAKSSVSSRSMSITSEASPNANLIGWKIGMLRIGECIKRMSTLIRLGSRAVIT 838
 Qy 651 KMFARKCQEAEANSSSGGYTORQNYLQQQQQDAPPNNPLITSPINSTLNLSUGTSTGG 710
 Db 839 KGPAR-----YNEQM-----849
 Qy 711 VEGGGINFGPNGPNTGNTNTGSPSPQQQQHQGTTNKSPPIIDPDRADIALYLG 770
 Db 850 -----EASPOQEIAMYLG 863
 Qy 771 IODHITTMQNLAYKTFSRSHSNYLQLOQVESFNSNNKITEMPSKILIGMVLPLNL 830
 Db 864 IQHIVTVASLSRHTKLSRSNLYLAQINIMTRVNLVLRKLTGIVLPMNV 923
 Qy 831 VTCLGFMNRYVPGEGCITLGNFGFIVGVLPIITGSPITPAQ 871
 Db 924 ITCGLMGMYVTFQGTYRDLSLTFGIVLFMCMLACASAYMTK 964
 RESULT 9
 Db Sequence 8, Application US/0894549
 ; Publication No. US20020138880A1
 ; General Information:
 ; FILE REFERENCE: 0894549
 ; CURRENT APPLICATION NUMBER: US/08/945/749
 ; CURRENT FILING DATE: 1998-01-12
 ; PUBLICATION NO.: US20020138880A1
 ; PCT/NZ96/0035
 ; GENERAL INFORMATION:
 ; FILE REFERENCE: 0894549
 ; CURRENT APPLICATION NUMBER: PCT/NZ96/0035
 ; CURRENT FILING DATE: 1998-05-01
 ; EARLIER APPLICATION NUMBER: NZ 22039
 ; EARLIER FILING DATE: 1995-05-01
 ; NUMBER OF SEQ ID NOS: 16
 ; SEQ ID NO: 8
 ; SOFTWARE: PatentIn Ver. 2.0
 ; LENGTH: 387
 ; TYPE: PRT
 ; ORGANISM: Synechocystis sp.
 ;
 RESULT 7
 Db Sequence 7, Application US/0894549
 ; Publication No. US20020138880A1
 ; General Information:
 ; FILE REFERENCE: 21402-012
 ; CURRENT APPLICATION NUMBER: US/09/864,408A
 ; Publication No. US20040009474A1
 ; PRIORITY INFORMATION:
 ; APPLICANT: Leach, Martin D.
 ; APPLICANT: Shinkles, Richard A.
 ; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encoded Thereby
 ; FILE REFERENCE: 21402-012
 ; CURRENT APPLICATION NUMBER: US/09/864,408A
 ; CURRENT FILING DATE: 2001-05-24
 ; PRIOR APPLICATION NUMBER: 60/206,690
 ; PRIOR FILING DATE: 2000-05-24
 ; NUMBER OF SEQ ID NOS: 9056
 ; SEQ ID NO: 2266
 ; LENGTH: 154
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens

US-08-945-749-8

Query Match 6.2%; Score 298; DB 2; Length 387;
 Best Local Similarity 23.0%; Pred. No. 4.8e-12;
 Matches 108; Conservative 70; Mismatches 143; Index 148; Gaps 18;

Qy 437 DRFSLPH---SESEETIHADIPS---IV--SPGOSTR--DLFENGEBTN----W 477
 Db 27 DYDFTYDEPGSEPTOLSLIEPAPPSTVILVDPSHARKSDISFLNRLPVLGNTVSW 86

Qy 478 LDCT-CPTDSKEMKLAFKGHPTAEDIRQETREKEVLFKYYTVCPTPDAKEDIS 536
 Db 87 MDIEGLGSEEVKEKEYGEFKKHPLLEDIVNPORAKEYDNDHVIAHYRPNRE-BD 145

Qy 537 YLEPINV---YIVVPHOGILTEHFSPISHANVRERVRQLRDYU-DVSADMILCYAL 588
 Db 146 GPESEGVSLVSKRULLPQSHLIDCNP---REIRITQKVCOQGADLYCL 199

Qy 589 IDEITDGAPVIGHGIEYADEIADEAFTERARTDEFSMLQRIGBSRKVATLRLSGRAD 648
 Db 200 TMLLDETFPLDEDTEENIAEDETTI--RNPN-SLMLMELYHIRELLALARLI---- 251

Qy 649 VIKMPAKRCQEANSSGTYQDNLQOOOOQAPPPNPNLITSPINSTNLNSLGTSNG 708
 Db 252 -----WPRHMYANWLDRD---- 265

Qy 709 GSVGYGGINFGNPGNTNTNTGSPSPPOOCQHGNTKNSFIDPARDPRADIALY 768
 Db 266 -----TNSIVT-----ADRVTRP 278

Qy 769 GDIQDHITMFONLLAYEKITSRSHSNTSNTLAQLOVESNSNKRTFSKLTIGMLDP 828
 Db 279 RDCYDHITQVLDITAYRELASLSMEYMTAM----SNKNEATKPLTWISTIFIP 331

Qy 829 NLVYGLPGRHNRVPGEGTNLUWPGFIVGULIFTIGSFIAOW---WL 874
 Db 332 TPIAGYGRANKFEMPE--LNSRGYIYTIVMLLIAAGSLYFFWRKGWL 378

RESULT 10
 Sequence 7005, Application US/11097143
 Publication No. US200502085581
 APPLICANT: Venter, J. Craig
 ATTORNEY: et al.
 TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
 TITLE OF INVENTION: DROSOPHILA GENES.

FILE REFERENCE: CL000728

CURRENT APPLICATION NUMBER: US/11/097,143
 CURRENT FILING DATE: 2005-04-04
 PRIOR APPLICATION NUMBER: 60/157,832
 PRIOR FILING DATE: 1999-10-05
 PRIOR APPLICATION NUMBER: 60/160,191
 PRIOR FILING DATE: 1999-10-19
 PRIOR APPLICATION NUMBER: 60/161,932
 PRIOR FILING DATE: 1999-10-28
 PRIOR APPLICATION NUMBER: 60/164,789
 PRIOR FILING DATE: 1999-11-12
 PRIOR APPLICATION NUMBER: 60/171,637
 PRIOR FILING DATE: 1999-12-28
 PRIOR APPLICATION NUMBER: 60/175,693
 PRIOR FILING DATE: 2000-01-12
 PRIOR APPLICATION NUMBER: 60/184,831
 PRIOR FILING DATE: 2000-02-24

SOFTWARE: PastSeq For Windows Version 4.0
 NUMBER OF SEQ ID NOS: 43008
 SEQ ID NO: 7905
 LENGTH: 1198

Query Match 4.9%; Score 234.5; DB 6; Length 1198;
 Best Local Similarity 19.7%; Pred. No. 5.7e-07; Mismatches 291; Index 349; Gaps 41;

TYPE: PRO^r
 ORGANISM: DROSOPHILA

US-11-097-143-7905

Query Match 4.9%; Score 234.5; DB 6; Length 1198;
 Best Local Similarity 19.7%; Pred. No. 5.7e-07; Mismatches 291; Index 349; Gaps 41;

RESULT 11
 US-08-945-749-7
 Sequence 7, Application US/08945749
 Publication No. US2008138880A1
 GENERAL INFORMATION:
 APPLICANT: GARDNER, Richard C
 ATTORNEY: MacDiarmid, Colin W
 APPLICANT: RAY, Robert J
 APPLICANT: Auckland Uniservices Limited
 APPLICANT: New Zealand Pastoral Agriculture Research Institut
 TITLE OF INVENTION: ALUMINUM RESISTANCE GENE
 FILE REFERENCE: 08/945-749
 CURRENT APPLICATION NUMBER: US/08/945,749

CURRENT FILING DATE: 1998-01-12
 ; EARLIER APPLICATION NUMBER: PCT/NZ96/0035
 ; EARLIER FILING DATE: 1996-05-01
 ; EARLIER APPLICATION NUMBER: NZ 272039
 ; EARLIER FILING DATE: 1995-05-01
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO: 7
 ; LENGTH: 380
 ; TYPE: PRT
 ; ORGANISM: Synochocystis sp.
 US-08-945-749-7

Query Match 86; Conservative 67; Mismatches 150; Indels 135; Gaps 13;

Qy 459 VSPGOSVADLPNGEE---TWWLDCPCPDSEKMLAKARGHPLTAEDIRQETREKV 514
 Db 62 VESPASLADTRELRDPSGAWGLARPESSLSLAEPFDLHPLAIVEDAEAHORPL 121

Qy 515 ELFKEYYYCFCHT--FEADKESEDYLEPINVYVPHDGLTPHFSPISSHANVRVHQ 572
 Db 122 ERYGETLUVLRAARYLDAPBEVDGF---LHFVGDPPVITRGAAPDLSAVARMBS 178

Qy 573 LRDYDVDSADMCLVALDETDGAPVIGIEYEADAEVATAVDTDESSNLQIGTS 632
 Db 179 TPELLKLGPPEAVLDAVDGTVPUAGVQNDIDEIEVEFRGD---PAVSRIYEL 234

Qy 633 BRKVNTLRLSLSKGADWKMPAKRCQEANSSGGYQROYNLQQQQQAPPNNPHTS 692
 Db 235 SREHTEVOP-----ATR 246

Qy 693 PINSTLNUNISLGSTCGGCVGGCINFGNPNNTNTGSPSPPOOCQHGJTKS 752
 Db 247 PLVGNLH-----SLNAGFAKYKTQ----- 265

Qy 753 FPIPDARPRADIALLYSDIQH----ITMPONLA----YEKIPSRSHSNLYLAQLQ 801
 Db 266 -----BELQRQLRDVAOYTHTSERVGGPQLADLTNTVYOOQNAERALA 316

Qy 802 VESENFSNKITEMSKTLMGTLMVPLNUTGJFGNVRVPGEGTNLGWFFG---IVGV 858
 Db 317 EAGFQNBEI---KKTSSMAMALEPFLVGTYGRNF---EHNDELGWSGYPAFTGL 368

Qy 859 LIPITGSFIP---RQWIL 874
 Db 369 MGLCVCSVLXVFKRGML 386

RESULT 13
 US-10-156-761-10644
 ; Sequence 10644, Application US/10/156761
 ; Publication No. US2003/0119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OKURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAI, YOSHIOUKI
 ; APPLICANT: HATORI, MASAHIRA
 ; FILE REFERENCE: 249-252
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; CURRENT FILING DATE: 2002-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO: 10644
 ; LENGTH: 371
 ; TYPE: PRT
 ; ORGANISM: Streptomyces avermitilis
 US-10-156-761-10644

Query Match 97; Score 201; DB 4; Length 371;
 Best Local Similarity 19.7%; Pred. No. 2.3e-05;
 Matches 91; Conservative 64; Mismatches 160; Indels 148; Gaps 13;

Qy 436 PDRESLPHSESE-ETIHAPIPSL-----SPQOSUPDLPNGCETW 476
 Db 13 PSRFSRKDKSATDTDPTPSAVVDFLHPLAIVEDAEAHORPL 532

Qy 477 WLDCTCPIDSMRMHLAKAGHPLTAEDIRQETREKVBFKSYFVCF--HTEPADK-
 Db 72 WIGHHEPEAFAEGLAGEPFHPLAIVEDAEAHORPL 531

Qy 533 -ESDYLEPINVYVPHDGLTPHFSPISSHANVRVHQDVSADWLCYALDE 591
 Db 132 TANSEVETEVGMCPDTPFVTRGGQSLSRALHRQDDEPLAEGPSAVLRA 191

Query Match 92; Conservative 213.5; DB 4; Length 386;
 Best Local Similarity 19.6%; Pred. No. 3.4e-06;

Db 192 VVDGTLAVADAVQDDIDEVETEVSPGRKTPRSTDAG---RIVQLKEVLFRKAVNSP 247
 Qy 646 KADVKMPAKRCQEANSSSCYXORINLQQQOQAAPPPEPNTITSPINSTLNLSCT 705
 Db 248 LLRPHQIISER-----PMFLID-----264
 Qy 706 STGGGVGCGTNGFPNPGNNTNTGSPSPQQQONGITNTSPSPDARPAIDA 765
 Db 265 -----PDIQ 268
 Qy 766 LYGDQDHITMFPNLAYERIKIFSRSHSNTYLAQLOVESNSNKITEMFSKITLIGTM 825
 Db 269 KFEDYDAHLLARVQEVQFLFEANSLQANLAQSAQNEQRKTS-WAIAIAVPTM- 326
 Qy 826 VPFLNLYTSGFLGMYRVPCEGGTNLGWFG---IVGVLIIFIG 865
 Db 327 -----VCGVYGRNF----EHPPELHWRFQGPVIMSVVWLUG 360

RESULT 14
 US-09-738-626-3568
 Sequence 3568, Application US/09738626
 Publication No. US20020197605A1

GENERAL INFORMATION:
 APPLICANT: NAKAGAWA, SATOSHI
 APPLICANT: MIZOGUCHI, HIROSHI
 APPLICANT: ANDO, SEIJI
 APPLICANT: HAYASHI, MIKIRO
 APPLICANT: OCHIAI, YUKIO
 APPLICANT: TOKOI, HARUHIKO
 APPLICANT: SENOH, AKIHIRO
 APPLICANT: IKEDA, MASATO
 APPLICANT: OZAKI, AOKI

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: Patentin ver. 3.0
SEQ ID NO: 3568
LENGTH: 373
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-3568

Query Match Score: 3.8%
Best Local Similarity: 19.9%
Matches 100: Conservative 69; Mismatches 144; Indels 189; Gaps 20;

Qy 432 GGATSDGINGSPLPDRFLSPHSESETHTAPDPLVSP-----GQSYR 466
 Db 9 GAIIRRDMLRPRYLPD-----SANSTPVPEEVSPTRVTFGPKPPYGAIV 57

Qy 467 D-LPRNGSEBTWLDCTCPDSE--RMDMLKAGIRPLTAEDIRMOTREKVLPKSTYF 522
 Db 58 DGLKFAEGASRMMMSLYPAPSATEELAEADLHPTIVEDLIGQRPKDRLIEDIF 117

Qy 523 VCFHP---PDAKESEDYLEPTNNYIVVPH-----DGILTHFEPSPISH 562
 Db 118 IAFSARYDREEDFSE-----FHLMKPQIAILCQDNQDGTGTSASF---SN 166

Qy 563 PANVRVRYQD--RDYDVDSAWMLCYALJDETDPSFAPVIGIEYTAEDDAYFTARD 619

Db 167 PEEDKRIKTTLADAELLSSGPRAVAYRLVDAVQDPSFSPVLAIDQBLEROF---222

Qy 620 TDFSSMQRICRSRKYMTHMLLSGKATVKAFCRQCBEEASSSSGYQORNWQQQ 679
 Db 221 -----SGDAAVA-----ERYNUSSQ-----238

Qy 680 QAPPPPEPNTITSPINSTLNLSCTGTSGVEGGINFGNPNTNNTNTGSSPSP 739

Db 239 -----IIM-----OHTTSVTP 251
 Qy 740 FOOCQHGITNKSPPIDARPADIALYLDIDHITMFQNLLAYEKIFSR---SHSNY 796
 Db 252 VVQR-----LAKKD--IRSGLSEEFRLRAYLDVADHLTRDNTVSEYRESLSQILNNTL 304
 Qy 797 LAQIQLQESPNSNNKITEMSKITLGMVPLNLYCPLGMYRVCCEGTTLGNWFG-- 854
 Db 305 VAQRO-----NEDMKRISWAIIAPFLVSSYGMNDIAP-----LHWAFGTP 350

RESULT 15
 US-11-082-389-248
 Sequence 248, Application US/11082389
 Publication No. US20050244935A1

GENERAL INFORMATION:
 APPLICANT: Pompejus, Markus
 APPLICANT: Kroger, Burkhard
 APPLICANT: Schroder, Hartwig
 APPLICANT: Zeider, Oskar
 APPLICANT: Haberhauser, Gregor
 APPLICANT: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
 CURRENT APPLICATION NUMBER: US/11/082,389
 CURRENT FILING DATE: 2005-03-16
 FILE REFERENCE: BCI-131CPCN
 TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MENERVANE
 PRIOR APPLICATION NUMBER: US/09/604024
 PRIOR FILING DATE: 2000-06-23
 PRIOR APPLICATION NUMBER: US/09/61031
 PRIOR FILING DATE: 1999-06-05
 PRIOR APPLICATION NUMBER: US/09/143262
 PRIOR FILING DATE: 1999-07-09
 PRIOR APPLICATION NUMBER: US/09/151281
 PRIOR FILING DATE: 1999-08-27
 PRIOR APPLICATION NUMBER: DE 19930487-4
 PRIOR FILING DATE: 1999-07-01
 PRIOR APPLICATION NUMBER: DE 19930489-0
 PRIOR APPLICATION NUMBER: DE 19931549.3
 PRIOR FILING DATE: 1999-07-08
 PRIOR APPLICATION NUMBER: DE 19931550-7
 PRIOR FILING DATE: 1999-07-08
 PRIOR APPLICATION NUMBER: DE 19932134-5
 PRIOR FILING DATE: 1999-07-09
 PRIOR APPLICATION NUMBER: DE 19941379-7
 PRIOR FILING DATE: 1999-08-11
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 446
 SEQ ID NO: 248
 LENGTH: 373
 TYPE: PRP
 ORGANISM: Corynebacterium glutamicum
 US-11-082-389-248

Query Match Score: 3.7%
Best Local Similarity: 19.9%
Matches 100: Conservative 66; Mismatches 145; Indels 189; Gaps 20;

Qy 422 GGATSDGINGSPLPDRFLSPHSESETHTAPDPLVSP-----GQSYR 466
 Db 9 GAIIRRDMLRPRYLPD-----SANSTPVPEEVSPTRVTFGPKPPYGAIV 57

Query Match Score: 3.7%
Best Local Similarity: 19.9%
Matches 100: Conservative 66; Mismatches 145; Indels 189; Gaps 20;

Qy 422 GGATSDGINGSPLPDRFLSPHSESETHTAPDPLVSP-----GQSYR 466
 Db 9 GAIIRRDMLRPRYLPD-----SANSTPVPEEVSPTRVTFGPKPPYGAIV 57

Qy 467 D-LPRNGSEBTWLDCTCPDSE--RMDMLKAGIRPLTAEDIRMOTREKVLPKSTYF 522
 Db 58 DGLKFAEGASRMMMSLYPAPSATEELAEADLHPTIVEDLIGQRPKDRLIEDIF 117

Qy 523 VCFHP---PDAKESEDYLEPTNNYIVVPH-----DGILTHFEPSPISH 562
 Db 118 IAFSARYDREEDFSE-----FHLMKPQIAILCQDNQDGTGTSASF---SN 166

Qy 563 PANVRVRYQD--RDYDVDSAWMLCYALJDETDPSFAPVIGIEYTAEDDAYFTARD 619

Db 167 PEEDKRIKTTLADAELLSSGPRAVAYRLVDAVQDPSFSPVLAIDQBLEROF---222

Qy 620 TDFSSMQRICRSRKYMTHMLLSGKATVKAFCRQCBEEASSSSGYQORNWQQQ 679

Db 221 -----SGDAAVA-----ERYNUSSQ-----238

Qy 680 QAPPPPEPNTITSPINSTLNLSCTGTSGVEGGINFGNPNTNNTNTGSSPSP 739

Db 167 PEEDKRIKTTLADAELLSSGPRAVAYRLVDAVQDPSFSPVLAIDQBLEROF---222

Qy	620	TDFSSMLQRGESEKRYVATMLRLLSGRADYVIRPKRCOPBEANSSGGYQYRONIQQQQQ	679
Db	223	-----SGDAVA-----	238
Qy	680	QAPPPIPNTSPINSTLNLSIUGSTGGVGVGINTGPNTGNNNTNTNTGNSPSP	739
Db	239	-----TIDM-----	251
Qy	740	PQQQHQHITGKSPIDARPADIALYDODHITMFQNLYLKEVIFSR-----SHRN 796	
Db	252	VVQR-----FRGSMSEBLAYLDVADHLTRNTVSERELSQLAVNATL 304	
Qy	797	LAQVQSPSNNKKITEMFKFLIGTMVPLAUYLTCGLGMWVPGCEGTTLGKPFPG-----854	
Db	305	VAQRQ-----NMDMKKISGWAIITAPTLVSIYGMFDIMP-----LHWAFGTP 350	
Qy	855	--IVGVLIPIIGSFIP--AQW 872	
Db	351	LALLAMGFTULLWIKRSKW 372	

ALIGNMENTS

SUMMARIES						
	Score	Match	Length	DB	ID	Description
1:	4820	100.0	922	3	AAB36515	Aab36515 Candida a
2:	4814	99.9	922	5	ABP73838	Abp73838 Candida a
3:	1679	34.8	859	2	AAM07872	Aam07872 Aluminium
4:	1666	34.6	858	2	AAM07873	Aam07873 Aluminium
5:	1099	22.8	663	6	ABJ25905	Abj25905 Aspergilll
6:	1099	22.8	663	6	ABJ26505	Abj26505 Aspergilll
7:	865	17.9	969	7	ADK62466	Adk62466 Disease t
8:	821	17.0	228	5	ABP05970	Abp05970 Human ORP
9:	345.5	7.2	154	5	ABP21160	Abp21160 Human q1Y
0:	319.5	6.6	107	5	ABP06753	Abp06753 Human ORP
1:	251	5.2	326	9	ABM92441	Abm92441 M. xanthum
2:	234.5	4.9	1198	4	ABB63371	Abb63371 Drosophil
3:	210.5	4.4	354	9	ABE0726	Aeb0726 L. pneumo
4:	210.5	4.4	357	9	AEB37409	Aeb37409 L. pneumo
5:	186.5	3.9	1072	5	ABBS4963	Abbs4963 Lactococc
6:	182.5	3.8	373	4	AAG9814	Aag9814 C. glutamini
7:	179.5	3.7	373	3	AED4911	Aed4911 Membrane
8:	179.5	3.7	1172	5	ABG93308	Abg93308 C. albica
9:	175	3.7	1199	4	ABB38274	Abb38274 Drosophil
0:	175	3.6	1246	4	ABB36157	Abb36157 Drosophil

卷之三

Result No.	Query		Length	DB	ID	Description
	Score	Match %				
1	4820	100.0	922	3	AAB36515	Aab36515 Candida a
2	4814	99.9	922	5	ABP73938	Aab73938 Candida a
3	1679	34.8	859	2	AAW07872	Aaw07872 Aluminium
4	1666	34.6	858	2	AAW07873	Aaw07873 Aluminium
5	1099	22.8	663	6	ABJ25905	Abj25905 Aspergill
6	1099	22.8	663	6	ABJ25055	Abj25055 Aspergill
7	865	17.9	969	7	ADK62466	Adk62466 Disease t
8	821	17.0	228	5	ABP05970	Abp05970 Human ORF
9	345.9	7.2	154	5	ABP31260	Abp31260 Human g1y
10	319.5	6.6	107	5	ABP07575	Abp07575 Human ORF
11	251	5.2	326	9	ABM22441	Abm22441 N. xanthu
12	234.5	4.9	1198	4	ABB60371	Abb60371 drosophil
13	210.5	4.4	354	9	AEB40726	Aeb40726 L. pneumo
14	210.5	4.4	357	9	ABT31409	Abt31409 L. pneumo
15	186.5	3.9	1072	5	ABP54963	Abp54963 Lactocac
16	182.5	3.8	373	4	AAG98814	Aag98814 C. glutami
17	179.5	3.7	373	9	ABD69971	Abd69971 Membrane
18	179.5	3.7	1172	5	ABG93308	Abg93308 C. albica
19	179	3.6	1199	4	ABB58274	Abb58274 Drosophil
20	175.5	3.6	1246	4	ABR61547	Abr61547 Drosophil

useful as drug targets. Fragments of them are useful as probes and primers for diagnosis of fungal infections, also as antisense and ribozyme agents. Proteins encoded by the genes are used to screen for their specific inhibitors which are potential antifungal agents for controlling a wide range of fungi pathogenic on animals or plants.

Fragments of the proteins are also used to raise specific antibodies.

Higher animals do not contain genes closely similar to the C. albicans genes, so antifungal agents that target them are less likely to be toxic to humans

XX

Sequence 922 AA:

Query Match	Score	DB	Length
Best Local Similarity	100.0%	4820;	922;
Matches	922;	Pred. No. 0;	
Conservative	0;	Mismatches	0;
Indels	0;	Gaps	0;
			:
1	MSDSESYNTSTNQPIRSDVELDHNRQNTMCATSDSEBELEKSLSELESEVKVSKQ	60	
1	MSDSESYNTSTNQPIRSDVELDHNRQNTMCATSDSEBELEKSLSELESEVKVSKQ	60	
61	QHHQBIBTSNAKPLTRKGSSKKKSENLTDKDRTTNPMLSGDDPDTNGKHANNNMS	120	
61	QHHQBIBTSNAKPLTRKGSSKKKSENLTDKDRTTNPMLSGDDPDTNGKHANNNMS	120	
121	SURKOFYKLKDTDNSTNTHLAIPIPPIPPIPITNANKSRKSQLENPLPLKKTKI	180	
121	SURKOFYKLKDTDNSTNTHLAIPIPPIPPIPITNANKSRKSQLENPLPLKKTKI	180	
181	GRNSARPFENDLVSPTAKTKNDSDTNTSTNTHMPLKLGIGATTLGVGTTATATA	240	
181	GRNSARPFENDLVSPTAKTKNDSDTNTSTNTHMPLKLGIGATTLGVGTTATATA	240	
241	AAGRPRPSRSIDSEADSHA SRSQSPTEDVCPMPVGDHIRVNGIDFDEIDEFIREBRA	300	
241	AAGRPRPSRSIDSEADSHA SRSQSPTEDVCPMPVGDHIRVNGIDFDEIDEFIREBRA	300	
301	YLQKQMTAKNTRIDEFONLSKAKNTTSGASRHPYTHHSNNNNKNGDGGSSMALKT	360	
301	YLQKQMTAKNTRIDEFONLSKAKNTTSGASRHPYTHHSNNNNKNGDGGSSMALKT	360	
361	PRNIKLUKTSPEFHENSSESETYLTKTQQPKYDQLSLTSTSTSGSGGVK	420	
361	PRNIKLUKTSPEFHENSSESETYLTKTQQPKYDQLSLTSTSTSGSGGVK	420	
421	FGGARISDINGSLPDRPLSPHSESEETIHAPIPSLVSPGQSVRLFRNGETWMDC	480	
421	FGGARISDINGSLPDRPLSPHSESEETIHAPIPSLVSPGQSVRLFRNGETWMDC	480	
481	TCPPTDSEKMKAFLAKGFIPLTAEDIHQETREKVLFKSTYFCFPTFADEKESEDYLEP	540	
481	TCPPTDSEKMKAFLAKGFIPLTAEDIHQETREKVLFKSTYFCFPTFADEKESEDYLEP	540	
541	INVYVPHDXGILTHSPKAVPARYTQLRDVDSADMCLVALDETDGPAPV	600	
541	INVYVPHDXGILTHSPKAVPARYTQLRDVDSADMCLVALDETDGPAPV	600	
601	RGEYBAIDBDAYPTARDTDSSENLRIGSRRKWTLMRLUSGADYIKMFARCB	660	
601	RGEYBAIDBDAYPTARDTDSSENLRIGSRRKWTLMRLUSGADYIKMFARCB	660	
661	ANSSGGYQRQYNLQOQQQAPPPPPNPTTSPINSTLNINSGTGTCGGVGGINFGP	720	
661	ANSSGGYQRQYNLQOQQQAPPPPPNPTTSPINSTLNINSGTGTCGGVGGINFGP	720	
721	NPGANTNTNTGSPSPQQQQHGTINKSPIDPARRADIALYLDIODHITMfq	780	
721	NPGANTNTNTGSPSPQQQQHGTINKSPIDPARRADIALYLDIODHITMfq	780	
781	NLLAYEKIFRSHSNYLAQLOVESFSNNNKITEMFSKILITGMVPLMLVTLGFMAYR	840	

Db 781 NLLAYEKIFRSHSNYLAQLOVESFSNNNKITEMFSKILITGMVPLMLVTLGFMAYR 840

Qy 841 VPGEGGTNLGMEFFGTVGVLPIIIGSPITPAOMWKLKANNSIECONNGNRPINHSSRSSI 900

Db 841 VPGEGGTNLGMEFFGTVGVLPIIIGSPITPAOMWKLKANNSIECONNGNRPINHSSRSSI 900

Qy 901 RSLGUKKKGNGKSTISIPNKYE 922

Db 901 RSLGUKKKGNGKSTISIPNKYE 922

RESULT 2

ABP73838

ID ABP73838 standard; protein: 922 AA.

XX

AC ABP73838;

XX

DT 30-JAN-2003 (first entry)

XX

DE Candida albicans essential protein SEQ ID NO 7675.

XX

KW Fungus; yeast; tetracyclin; promoter; GRACE strain; biosynthesis;

KW signal transduction; DNA replication; cell division; growth;

KW proliferation; Candida albicans; fungicide; antifungal.

OS Candida albicans.

XX

PN WO200253728-A2.

XX

PD 11-JUL-2002.

XX

PP 26-DEC-2001; 2001WO-US049486.

XX

PR 29-DEC-2000; 2000US-0259128P.

PR 20-FEB-2001; 2001US-00792024.

PR 22-AUG-2001; 2001US-0314050P.

XX

PA (ELT-T) ELTTRA PHARM INC.

XX

PI Roemer T., Jiang B., Boone C., Bussey H., Ohlseni KL;

XX

DR WPI: 2002-565694/60.

DR N-PSDB: AB32388.

XX

PT Constructing strains for identifying gene products as effective targets for therapeutic intervention, by inactivating in the strain one allele of a gene and placing other allele of the gene under conditional expression.

XX

PT Claim 44: SEQ ID NO 7675; 167pp + Sequence Listing: English.

XX

CC The invention relates to constructing (M1) a strain of diploid fungal cells in which both alleles of a gene are modified, comprising modifying one allele by insertion or replacement by cassette having an expressible selectable marker and modifying other allele by recombination, of a promoter replacement fragment with a heterologous promoter, that expression of the second allele is regulated by the promoter. (M1) is useful for constructing a strain of diploid fungal cells in which both alleles of a gene are modified. The diploid fungal cells having both alleles modified are useful for identifying a gene that is essential to the survival or growth of a fungus, a gene that contributes to the virulence and/or pathogenicity of a fungus, a gene that contributes to the resistance of a diploid fungus to an antifungal agent, an antifungal agent that inhibits the growth of diploid fungus and for identifying a therapeutic agent for treatment of a mammalian disease. (M1) is useful for identifying a compound which modulates the activity of a gene product, preferably enzymatic activity, carbon compound catabolism, biosynthesis, transcription, translation, signal transduction, DNA replication and cell division activity. The method is useful for identifying a compound having the ability to inhibit growth or proliferation of C. albicans cells and for

treating infection by *C. albicans*. The present sequence is that of an essential *Candida albicans* protein used in the method of the invention.

Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office

Sequence 922 AA:

Query Match 99.9%; Score 4814; DB 5; Length 922;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 921; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 841 VPGEGTNLGNFFGIVSVLFLIIGSPFIAQWVLLKLNNSIECQNNNGRPIFNHSSRSI 900
 Qy 901 RSLGKXKGKNSKTSISPPNTE 922
 Db 901 RSLGKXKGKNSKTSISPPNTE 922

SQ Sequence 922 AA:

RESULT 3
 AAW07872 ID AAW07872 standard; protein: 859 AA.
 XX ID XX
 AC AAW07872;
 DT 23-FEB-1997 (first entry)
 DE Aluminum resistance gene ALR1.
 KW Aluminum resistance; ALR1; ALR2; ARH1; cation; tolerance;
 KW magnesium transport.
 XX Sacccharomyces cerevisiae.
 OS OS XX
 PN W0963495-A1.
 XX XX
 PD 07-NOV-1996.
 XX XX
 PF 01-MAY-1996; 96WO-NZ000035.
 PR PR 01-MAY-1995; 95NZ-0072039.
 PA PA (AUCK-) AUCKLAND UNISERVICES LTD.
 PI Gardner RC, Macdiarmid CM, Hay RJM;
 DR DR WPI: 1996-506161/50.
 XX XX
 PT AHRI-homologues, aluminium resistance genes ALR1 and ALR2 - used in the treatment of cation deficiency, cation toxicity, esp. heart disease.
 PS PS Claim 6; Fig 5; 62pp; English.
 XX XX
 CC ALR1 and ALR2 are isolated by selecting vectors which confer high aluminium tolerance, or that complement knockout mutations in ALR1, ALR2 and/or ARH1. The genes have homology with bacterial genes responsible for divalent ion uptake. The ALR1 and ALR2 are magnesium transport genes. The isolated transport genes can be used in the treatment of any plant, animal or microorganism disease which results from a cation deficiency, by producing an accumulation of cations in a plant deficient to those cations, or in plants consumed by animals deficient in those cations. The genes may also be used in the treatment of diseases resulting from cation toxicity, esp. manganese toxicity and esp. for the treatment of heart disease. Note: The ALR1 and ALR2 gene sequences are claimed and stated as given in Figure 6 and Figure 7 of the specification, however these figures are missing from the specification

PS Sequence 859 AA;

Query Match 34.8%; Score 1679; DB 2; Length 859;
 Best Local Similarity 41.5%; Pred. No. 6; 2e-19;
 Matches 405; Conservative 137; Mismatches 258; Indels 176; Gaps 31;

Qy 1 MSDSESYTQNNTNPQIPRSDEVDDHRNOTINDCATSDSEDELEKSRLESEEVTKSKQ 60
 Db 61 CQHQHQTDSNAKPLTRKSGSIIKKSNLTDKDRITNPMSLSGGDTTNSGHKNNYNTMS 120
 Qy 61 CQHQHQTDSNAKPLTRKSGSIIKKSNLTDKDRITNPMSLSGGDTTNSGHKNNYNTMS 120
 Db 121 SIRKOFYIKDNTDNSTNNNETHLAIPPIPIPITANKNSRKERKQLENLPLIKKTII 180
 Db 121 SIRKOFYIKDNTDNSTNNNETHLAIPPIPIPITANKNSRKERKQLENLPLIKKTII 180
 Qy 181 GRNNNTENDLUSPATKATNQSEDINTNSTANRKLIGATISVOTGTATATA 240
 Db 181 GRNNNTENDLUSPATKATNQSEDINTNSTANRKLIGATISVOTGTATATA 240
 Qy 241 AGRRPSSSIDSEADSHASPSQETEDVCPMPGDHIVNGIDDEDEFTRBEEA 300
 Db 241 AGRRPSSSIDSEADSHASPSQETEDVCPMPGDHIVNGIDDEDEFTRBEEA 300
 Qy 301 YLQKOMIARNTLIRDEFONLSKNTTSGASRHYPYHHSNNNNKNNNGDGOGGSSMALKYT 360
 Db 301 YLQKOMIARNTLIRDEFONLSKNTTSGASRHYPYHHSNNNNKNNNGDGOGGSSMALKYT 360
 Qy 361 PKNLKLTLSPREPHTNSSESEBILKTKQOPPKYDQQLSLTSSTSSTSGSGGVK 420
 Db 361 PKNLKLTLSPREPHTNSSESEBILKTKQOPPKYDQQLSLTSSTSSTSGSGGVK 420
 Qy 421 PGGATISDINGSPDRPSLPHESSEETHADPSLSPGQSVDFPRNGEWTWDC 480
 Db 421 PGGATISDINGSPDRPSLPHESSEETHADPSLSPGQSVDFPRNGEWTWDC 480
 Qy 481 TGPDTSEMNMKLAKGTHPLTAEDIHQETREKVELPKSKVYPCPHTFAKESEDYLYEP 540
 Db 481 TGPDTSEMNMKLAKGTHPLTAEDIHQETREKVELPKSKVYPCPHTFAKESEDYLYEP 540
 Qy 541 INVYIVPHDGLTHPSLPHSPISHPANVRRTQRLDYVDASADMCLYALIDEITDGPAPVII 600
 Db 541 INVYIVPHDGLTHPSLPHSPISHPANVRRTQRLDYVDASADMCLYALIDEITDGPAPVII 600
 Qy 601 HQIETAAIDIAVFTADTFSSMLQIGESRKRTMLRSLGADYTKMFKRCQEE 660
 Db 601 HQIETAAIDIAVFTADTFSSMLQIGESRKRTMLRSLGADYTKMFKRCQEE 660
 Qy 661 ANSSGGYQRQYNLQQQQQAAPPPIPPIITSPIASNLANSAGTGGCVGFGNFGP 720
 Db 661 ANSSGGYQRQYNLQQQQQAAPPPIPPIITSPIASNLANSAGTGGCVGFGNFGP 720
 Qy 721 NPGGNNTNTNTGSPSPPOOCQHGTTMKSPPIPDPARPADIALYLGDIQHITMFQ 780
 Db 721 NPGGNNTNTNTGSPSPPOOCQHGTTMKSPPIPDPARPADIALYLGDIQHITMFQ 780
 Qy 781 NLLAYEKIPSRSNLYLAQLOVESFNSNNKITEMFSKILIGTMVPLNLVTCGFMNVR 840
 Db 781 NLLAYEKIPSRSNLYLAQLOVESFNSNNKITEMFSKILIGTMVPLNLVTCGFMNVR 840
 Qy 841 VPGEGTNLGNFFGIVSVLFLIIGSPFIAQWVLLKLNNSIECQNNNGRPIFNHSSRSI 900
 Qy 913 KNR-----NYMSSL-----RKDFYLKDNTDNSTNHT-----HLAIPPIPPTII 156

Oy	334 YHHHSANDKCKNGDGGSSMAALKTPKNILKKTLSRPBPTHENSSSBEEYELKTQO	393	PR 31-AUG-2001; 2001US-0316362P.
Db	319 YHPTQDGTSFTNSASTGSS-AAKYTP-E---USR---TUEKNCSTNEYXENNES	368	XX XX PA (ELITR) ELITRA PHARM INC.
Oy	394 PPKYKDQLSLTSSTSSTSGSGSQQVREGARTSDFGNG---	437	PI Jiang B., Tishkoff D., Zamudio C., Broshkin AM., Hu W., Lemieux SM.
Db	369 --VREDDKPLD-----HPDVTGPNKLEGKEGRNDSSYRAYTTLQNTEQIPS	415	XX XX DR WPI; 2003-093124/08.
Oy	438 RFLSFHSESETHADPISLVSFGQSVDLFRNGEBTWLDCTCPDSEMMLAKAFGI	497	XX New purified or isolated nucleic acids of essential genes of Aspergillus fumigatus, useful for treating or preventing infections by A. fumigatus.
Db	416 RFSFERSBSEDEWIASDPSLSSCOPYBFLKGDPPTWMLDSCPDDEMRCIAKTFGI	475	XX PT PT or for treating a non-infectious disease in a subject e.g. cancer.
Oy	498 HPLTAEDIRMQTREKVELPKSYTFVCFCHTEFAKEDSEYDLEPINIVTVFDHGILTHP	557	XX PS Disclosure: Page: 175pp; English.
Db	476 HPLTAEDIRMQTREKVELPKSYTFVCFCHTEFAKEDSEYDLEPINIVTVFDHGILTHP	535	XX CC The invention relates to novel purified or isolated nucleic acids of essential genes of Aspergillus fumigatus. The isolated nucleic acids of the invention are used to treat or prevent infections by a pathogenic organism such as A. fumigatus, to treat a non-infectious disease in a subject (e.g. cancer), to prevent or contain contamination of an object by A. fumigatus, or to prevent or inhibit formation on a surface of a biofilm comprising A. fumigatus. The polynucleotides are useful for expressing recombinant protein for characterisation, screening or therapeutic use, as markers for host tissues in which the pathogenic organisms invade or reside, for comparing with the DNA sequence of A. fumigatus to identify duplicated genes or paralogues having the same or similar biochemical activity and/or function, for comparing with DNA sequences of other related or distant pathogenic organisms to identify potential orthologous essential or virulence genes, for selecting and making oligomers for attachment to a nucleic acid array for examination of expression patterns, for raising anti-protein antibodies, as an antigen to raise anti-DNA antibodies or to elicit another immune response, and for identifying polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction. The polypeptides may be used to raise antibodies or to elicit immune response, as a reagent in assays designed to quantitatively determine levels of the protein in biological fluids, as a marker for host tissues in which pathogenic organism invade or reside, and to isolate correlative receptors or ligands in the case of virulence factors. This sequence represents a protein of one of the essential genes of Aspergillus fumigatus of the invention
Oy	558 SPISHANVRVRVRLRDYDVSADMCLAYALIDBTGPAVINGTYEADAEVAFITA	617	XX CC Sequence 663 AA;
Db	536 DPISHANVRVRVRLRDYDVSADMCLAYALIDBTGPAVINGTYEADAEVAFITA	595	CC CC Best Local Similarity 32.7%; Prod. No. 9.7e-75; Mismatches 96; Indels 262; Gaps 22; Matches 292; Conservative 96; Mismatches 241; Indels 262; Gaps 22;
Oy	618 RNDPSSMLQRIGRSRRTVMTLMLRSLGRADVIMPAKRCRQDITGPAVINGTYEADAEVAFITA	674	CC CC 4 SEYYQNNTNQIPRPSDEVLDHNRQITNDCAISDDBLELKSEBSLEYVKSEKOOOH 63
Db	596 RMDPAAHLQRIGRSRRTVMTLMLRSLGRADVIMPAKRCRQDITGPAVINGTYEADAEVAFITA	655	CC CC 7 SEKTMCLTRPSTPQE---LDHHRQL---DSPPRIATADI-SLSRQNTQHAY 54
Oy	675 QQQQQAPPPNPNTSPINSTNLNLISLTSTGGVYGVGTMGPNPQMNNTNTNTT	734	CC CC 64 HOETITSDNAKPLTRKGSISKKSNLTDKDRITNPMSLSGGDNTTNSGHKNTYAMSSLR 123
Db	656 QAEQ---NVRQ-----NNSN-----	668	CC CC 55 HQEPQO-----NTRDNTSNH-----THLAIPITPIPTITMANKSRRK--SQ 167
Oy	735 GPSPPQQQDQGTTNKSPPDAPRADALVYDQDQHITMQLLAVKETKSRSHS	794	CC CC 124 KDFYLKD-----NTRDNTSNH-----THLAIPITPIPTITMANKSRRK--SQ 167
Db	669 -----NQISLN-SYMCYTSQPGRDIALVYDQDQHLMTEFQLLAVKETKSRSHS	718	CC CC 88 DDRSGKDINTLGRRFSVDPNGVRHGRTWSRTHQEL-----ANMSRESSPA 134
Oy	795 NYLAQLOQEVESPNINNKITEMSKTLMVLPUMLVNLTVGGRVPOEGGTNLGWFFG	854	CC CC 168 LENUPPLIKKTTGRNNSNNPENDLUSPTKMKTNSEDITNTSTANHMLIGARTLG 227
Db	719 NYLAQLOQEVESPNINNKITEMLGKTMGKTMGITHMLVPLVNTGICGMNVVPGRNG-STAWNYG	777	CC CC 135 RSSPP-----NSVEAFDPRRERANTL-----SHAAPDLEAIIQRTVSG 176
Oy	855 TVGVLIPIITGGSPTPAQWKLKLNST---ECONGNRP-----PNHSSRSI 900		CC CC 228 VGTST-TATATATAAGRRPSRSIDEADSHASRSQTEEDVCPMVGDHIVNGIDP 286
Db	778 ILGVLLLAWSWPLASWIKKIDPPTATLNEAAGSKAVISSPLKDKRFNDSSKGN	837	CC CC 177 ---GTHPRPFTFSNASAIRQPCDQILEPN-----DESCVPTYEQGRIPVTDY 222
Oy	901 RSLGJUKKHGKRSITISPKNY 921		XX XX DB 135 RSSPP-----NSVEAFDPRRERANTL-----SHAAPDLEAIIQRTVSG 176
Db	838 ARVGYRR---KSTVSLPSRY	854	XX XX DB 124 KDFYLKD-----NTRDNTSNH-----THLAIPITPIPTITMANKSRRK--SQ 167
RESULT 5			
ID	ABJ25905	standard: protein: 663 AA.	Qy 168 LENUPPLIKKTTGRNNSNNPENDLUSPTKMKTNSEDITNTSTANHMLIGARTLG 227
XX			Db 135 RSSPP-----NSVEAFDPRRERANTL-----SHAAPDLEAIIQRTVSG 176
AC	ABJ25905;		Qy 124 KDFYLKD-----NTRDNTSNH-----THLAIPITPIPTITMANKSRRK--SQ 167
XX			Db 88 DDRSGKDINTLGRRFSVDPNGVRHGRTWSRTHQEL-----ANMSRESSPA 134
DT	16-APR-2003	(first entry)	Qy 228 VGTST-TATATATAAGRRPSRSIDEADSHASRSQTEEDVCPMVGDHIVNGIDP 286
XX			OS Aspergillus fumigatus.
Aspergillus fumigatus essential gene protein #563.			XX XX
XX			PN WO20028090-A2.
KW Fungicide; cytosatic; essential gene; Aspergillus fumigatus; infection; cancer; contamination; biofilm; antibody; immune response.			XX XX PD 31-OCT-2002.
OS Aspergillus fumigatus.			PP 23-APR-2002; 2002WO-US013142.
XX			XX XX PN 23-APR-2001; 2001US-0285697P.
PR 27-APR-2001; 2001US-0287061P.			PR 05-JUN-2001; 2001US-025830P.
PR 05-JUL-2001; 2001US-030389P.			PR 223 EELSEVFTA-----LSRQMPSTSR-----KQSLSSQSGPVPYDLCRCLKS 266

Oy	344 NNGDGGSMAALKYTPKNTLKKTLRFBTHESSSETEYLTKQQPPYKYDOLIS 403	XX	New purified or isolated nucleic acids of essential genes of Aspergillus fumigatus, useful for treating or preventing infections by A. fumigatus, or for treating a non-infectious disease in a subject e.g. cancer.
Db	267 DVEGERKSSA-----DRSSDLMDALITADKTYANTVDEKD 304	PT	
Oy	404 LTSSTSSTSGSGQVKFGGARISQDINGSLPPIFSEETTHADPLSPVSPSQ 463	PT	
Db	305 -----IVEKLQNEPTEPRGFSSSEQSOSTVHAELGDLVLPEQ 342	PT	
Oy	464 SYRDIFRNGE--TMIWLDCTCPDTSENKMLAKAFGTHPIAEDTRAQETREKVELPKSY 521	PS	Disclosure; Page: 175pp; English.
Db	343 TFRDIFQLQEGWVWLVLPATPBEYAAALSRAFSPHTPLDILTEAREKVELPKSY 402	CC	The invention relates to novel purified or isolated nucleic acids of essential genes of Aspergillus fumigatus. The isolated nucleic acids of the invention are used to treat or prevent infections by a pathogenic organism such as A. fumigatus, to treat a non-infectious disease in a subject (e.g. cancer), to prevent or contain contamination of an object by A. fumigatus, or to prevent or inhibit formation on a surface of a biofilm comprising A. fumigatus. The polynucleotides are useful for expressing recombinant protein for characterisation, screening or therapeutic use, as markers for host tissues in which the pathogenic organisms invade or reside, for comparing with the DNA sequence of A. fumigatus to identify duplicated genes or paralogues having the same or similar biochemical activity and/or function, for comparing with DNA sequences of other related or distant pathogenic organisms to identify potential orthologous essential or virulence genes, for selecting and making oligomers for attachment to a nucleic acid array for examination of expression patterns, for raising anti-protein antibodies, as an antigen to raise anti-DNA antibodies or to elicit another immune response, and/or identifying polynucleotides encoding the other protein which binds occurs or to identify inhibitors of the binding interaction. The polypeptides may be used to raise antibodies or to elicit immune response, as a reagent in assays designed to quantitatively determine levels of the protein in biological fluids, as a marker for host tissues in which pathogenic organism invade or reside, and to isolate correlative receptors or ligands in the case or virulence factors. This sequence represents a protein of one of the essential genes of Aspergillus fumigatus of the invention
Oy	522 PFCPHTP-BADKESEDVLEPINVVIVPHGILTFPHSPISHDANVRVRLQDVKYVS 580	CC	
Db	403 FVCFRTFQIQLKTDERMEPKPNFTMVYFRQDVLISFSTENPHANTRKIGLRLDVLS 462	CC	
Oy	581 ADWLYVALIDEIOTDOPAVINGTEVEDAEDAWVTPARDTDSSMQRGEBSRRYMTM 640	CC	
Db	463 SWICVAMIDDVDSFGVTEIRETIESEATEDLWPLARVDDFESLPRIGGLRKVXMSLM 522	CC	
Oy	641 RLLSGRADVTHPAKRCOEBAANSSGGTYQORYNQ000000APPBPNPITSPISTNLN 700	CC	
Db	523 RLIGKADTVTRGSKRCNE----- 541	CC	
Oy	701 NSLGFTSTGGCQVGCGINFGPNTGPNNTNTNTGSPSPPQQQQHGTINKSFPIDARP 760	CC	
Db	542 -----QSVT----- 547	CC	
Oy	761 RADIALYLGQDHDITIMPONLAYEKTSRSHTNLAOLQVESNSNNKITEMFSKTT 820	CC	
Db	548 RDIGIYLGQDHDVVTMSSLARFKMLSRSHNTNLAQUNVNTLQGHVNKVLSKVT 607	CC	
Oy	821 IGTMLAVPLNLYTGFLGKDNVRYPGEGGCTNLLGWEFGTNGVLFITIGSFIAOM 873	CC	
Db	608 IATMLVPMNLTCGLFGMNVRYPGEGQSGCAGTGFGVIAVIVLSGIARY 660	CC	
RESULT 6			
ID	ABJ26505 standard, protein: 663 AA.	Matches	292; Conservative 98; Mismatches 241; Deletions 262; Gaps 22;
XX		Query Match	22.8%; Score 1099; DB 6; Length 663;
AC		Best Local Similarity	32.7%; Pred. No. 9.7e-5;
XX		Matches	292; Conservative 98; Mismatches 241; Deletions 262; Gaps 22;
DR	16-APR-2003 (first entry)	Sequence 663 AA;	
XX	Aspergillus fumigatus essential gene protein #1163.		
XX	W0200286090-A2.		
KW	Pesticide; cycostatic; essential gene; Aspergillus fumigatus; infection; cancer; contamination; biotin; antibody; immune response.		
XX			
ED	31-OCT-2002.		
XX	Aspergillus fumigatus.		
EN	WO200286090-A2.		
XX			
ED	31-OCT-2002.		
XX	Aspergillus fumigatus.		
EP	23-APR-2002; 2002WO-US013142.		
XX			
FR	23-APR-2001; 2001US-0285697P.		
FR	27-APR-2001; 2001US-0287065P.		
FR	05-JUN-2001; 2001US-0295890P.		
FR	09-JUL-2001; 2001US-0303899P.		
FR	31-AUG-2001; 2001US-0316362P.		
XX			
PA	(ELIT-) ELTRA PHARM INC.		
XX			
FI	Jiang B, Tishkoff D, Zamudio C, Broshkin AM, Hu W, Lemieux SM;		
XX			
DR	WPI: 2003-093124/08.		

Db	305	-----TVEKLONENEPPTPREGPPSESSESOSTVHAELCDLVLPED	342	CC second protein, or its derivative, fragment, homologue or variant. The CC proteins are selected from given protein complexes, which are not defined in the specification. The variants are encoded by nucleic acids that CC hybridize to the nucleic acids encoding the proteins under low stringency CC conditions. The protein complexes are useful as targets for an active agent of a pharmaceutical. These protein complexes are particularly useful as drug targets for the treatment or preventing of a disease or disorder. The complexes and methods above are useful in diagnosing or screening for the presence of a disease or disorder or a predisposition for developing a disease or disorder in a subject. These are also useful in screening for a drug for treatment or prevention of a disease or disorder. The molecule that modulates the amount, activity or protein components of the complex is useful for the manufacture of a medicament for the treatment or prevention of a disease or disorder. This sequence corresponds to a protein of the invention. (Note: the sequence data for this patent did not form part of the printed specification but was obtained from the EPO in electronic format).
Qy	464	SVRDLFRNGEE--TWLDCCTPTDSEKMLAKAGIPLAEDIRQETREKVELYSY	521	XX
Db	343	TFDLQLQBLGQVWVLQDNTEEVAVLSRASIRHLTDILTBAREKVLFQRY	402	XX
Qy	522	FVCFHTP--BAAKESDYLEPLTEPVNIVVPHDGLTHPSPISHPANYVRVRLRDYDV	580	CC
Db	403	FVCFHTP--BAAKESDYLEPLTEPVNIVVPHDGLTHPSPISHPANYVRVRLRDYDV	462	CC
Qy	581	ADWCVVALIDETDGPAVTHGIEVAEATEDAYPTARDTDESSMLQRIGESRRKMMUTM	640	CC
Db	463	SDWCYTAMIDIVDSFGPVIREBIESAIDLFLIAVRVDDFESFSPRIGLRKKYMSM	522	CC
Qy	641	RLLSGRADYVKAFAKRCOPEBEANSSGGTQRYQYNQQQQQAPPNNPNTSINSTNL	700	CC
Db	523	RLLSGRADYVKAFAKRCOPEBEANSSGGTQRYQYNQQQQQAPPNNPNTSINSTNL	541	CC
Qy	701	NSLGSTTGCGVGUGG INF GPNPTGANTNTNTGSPSP0000QHITNKSPIDARP	760	XX
Db	542	-----P 547	Sequence 969 AA:	XX
Qy	761	RADIALYLGDIODHHTMFQNLLAYERKTFSRSHSYNLAQLOVESFSNNKITEFSKTTL	820	Query Match
Db	548	RGDIGLYLSDTQDHVTTMNSLAIPERKMLSLTSITVNLULGHNVNKVLSKVTI	607	Best Local Similarity 24.2%; Fred. No. 1.1e-50;
Qy	821	I GTMVLPLMLVTLGFMGNVRYVPGEGTNLGRFEGVIFLITGSPFAQWN	873	Matches 271; Conservative 143; Mismatches 247; Indels 460; Gaps 29;
Db	608	IATMVLPAALICLGFGNVRYVPGEGTNLGRFEGVIFLITGSPFAQWN	660	Qy 11 SNTNPIRSPDEVLDLHRNQNTNDCAISDSEDLELKSELESEVERSE-----58
Qy	7	ADK62466 standard; protein: 969 AA.	Db 44 SDSRPTQPLHDNQLQHNGQTDEDQI-DSWGMAHESDTSNDIKSEDPSLKGAFIDHR 102	
AC	ADK62466;	XX	Qy 59 ----KQOQHQBITSDS-NARFLTRKSSSIIKKSNLNDKRTINPMS-USGGDDTINSGH 112	
XX	06-MAY-2004 (first entry)	XX	Db 103 PMSQPRGPQSVSSTVQPQPKKFSTPSYTKKAGLRSQNSVLSVDSPSE--LESWL 160	
DB	Disease treating protein complex-derived protein #359.	XX	Qy 113 KNRNTNMSLKDLYLKDNTDDNISTNRPHTHALAPIPIPTIITNA----NRSRK 165	
RESULT 7	XX	XX	Db 161 KER---KSVHKSPVDENSPTDPRQSHANNDVVYD-VLMHNVNASTGYNDSKRK 215	
ID	ADK62466	XX	Qy 166 SOLENLPPLIKKKTIGRNNNNPNDLUSPMTMKNDSEDTNTTSTANHMKLGIGATT 225	
XX	OS Unidentified.	XX	Db 216 -----KXR--GSDSSN-KN--SKSTSDDNDEEDNTS-----244	
XX	PA EP1338608-A2.	XX	Qy 226 LGVGTGTTATAAAGRPRESIDEADSASHASSQTEEDVCFPM--YGDITRYNG 283	
XX	PA 27-AUG-2003.	XX	Db 245 -----RESSS-----LSNNSSLLDQDVLLVLDGESETPKAW 275	
XX	PA 20-DEC-2002; 2002EP-00102902.	XX	Qy 284 IDPFDIDEPTREREAYLQKQMTAKNMRIDFQNLSKNTNTGASHRHPYBHHSNNKK 343	
XX	PA 20-DEC-2001; 2001EP-00130253.	XX	Db 276 PDCTVLEPSTEYER - LRS QAI-----ODAFAFOYDDE 312	
XX	PA (CBL-) CELZONE AG.	XX	Qy 344 NNGGGGGGSSMAALKTPKNNLKKLTSRPFETHENSSEYYELKTKQPPYKDQLS 403	
XX	PA WPI: 2003-638460/61.	XX	Db 313 D-----GTSNDGTLFSKPKTNTDPELGNRVTNETENLKGRLPKRIAWHJLQIPM 367	
XX	DR N-PSDB; AD62467.	XX	Qy 404 LTSSSTSSSGSSQGVFKGARIEDGINGSL-----PD-RPSLHSES 446	
XX	Bauer A, Superti-Furga G, Kuester B, Schultz J;	XX	Db 368 VL-----GSNSTKDSRSTQSGIQLNLLVGRNIQYPHIIISNNPHEFRPTYRVL 418	
FI	Marzioch M, Grandi P, Krause R, Kruse U, Merino A, Bauch A;	XX	Qy 447 EETHADPPIPUSPGPSVDF-----469	
FI	Michon A, Lentwein C, Rick J;	XX	Db 419 DSTVHSPTSCUOLPGKQFDPDLYVASTYQDNEAGHIKTHNSPTGIGAKETVYQGQT	
XX	DR	XX	Qy 470 -----RNGEET--WLDCTCPDSEKMLAKAFGTHPLTAEDIRNQETRYKV 515	
XX	New proteins and protein complexes from eukaryotes, useful as targets in	XX	Db 479 AKNPSTLSSMSVANTEIDVPPFWLDSNPIEEKVKLSKAFGTHPLTTEIFGEYREV 538	
PR	drug screening, or in diagnosing or screening for the presence of a	XX	Qy 516 LPKSYFFVCPHTFB-----ADKB -----534	
PR	disease or disorder, or a predisposition for developing a disease or disorder in a subject.	XX	Db 539 LFQDYLICLPSDIVAEKHMRRKEQESATLDEHESLSRKSKQAYGATMSNESNANN 598	
PS	Disclosure: SEQ ID NO 717; 13pp; English.	XX	Qy 535 -----535 -----	
CC	The invention relates to novel protein complexes comprising a first and a	XX		

PT	hyperproliferative disorders and autoimmune disorders.
XX	
PS	
Disclosure:	SEQ ID NO 11922; 1037pp; English.
XX	
CC	The present invention describes substantially purified human proteins referred to as open reading frame, ORPX, where X is 1-11491 (see Table 1 in the specification). ABN15762 to ABN27352 encode the human ORPX proteins given in ABP0001 to ABP1500. ORPX proteins are useful for treating or preventing a pathology associated with an ORPX-associated disorder in humans, and in the manufacture of a medicament for treating a syndrome associated with ORPX-associated disorder. ORPX polynucleotide sequences can be used in gene therapy. ORPX sequences can be used in the treatment of cancer, hyperproliferative disorders, cirrhosis of liver, psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage, osteoarthritis, neurodegenerative disorders, disorders related to organ transplantation, cardiovascular diseases, diabetes mellitus, systemic lupus erythematosus, hypertension, hypothyroidism, cholesterol ester storage disease, various immune deficiencies and disorders, infectious diseases, autoimmune disorders such as multiple sclerosis, rheumatoid arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host disease and autoimmune eye disease. ORPX proteins are also useful for treating burns, incisions, ulcers, for treating osteoporosis, bone degenerative disorders, or periodontal disease, and for gut protection or regeneration, or treatment of lung or liver fibrosis, reperfusion injury in various tissues and conditions resulting from systemic cytokine damage. N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 228 AA:
Query	Match Score 821; DB 5; Length 228;
Best Local Similarity 59.6%; Pred. No. 3.8e-54;	Gaps 5;
Mismatches 23; Conservative Matches 170;	Indels 60;
Db	1 TREKVLEKSTKYYFICPRHPPRKEADKSEDYLSPIVNYIVPHDGLMPHPSPISHPANVRR 569
Qy	510 TREKVLEKSTKYYFICPRHPPRKEADKSEDYLSPIVNYIVPHDGLMPHPSPISHPANVRR 569
Db	1 TREKVLEKSTKYYFICPRHPPRKEADKSEDYLSPIVNYIVPHDGLMPHPSPISHPANVRR 60
Qy	570 VRQLRNTVDYSDMLCYALIDDTGAPHTKGLEYTAIDENAYTARDTDPSEMMLQRI 629
Db	61 VRQLRNTVDYSDMLCYALIDDTGAPHTKGLEYTAIDENAYTARDTDPSEMMLQRI 120
Qy	630 GSSRKRTMVLRLSGRADYTKPFAKICQBEANSSCCYQRQ---YNLQOQQDQAPPDP 686
Db	121 GSSRKRTMVLRLSGRADYTKPFAKICQDETGIGPVLKSGQAE----- 173
Qy	687 NPLITSPINSTLNLSLGTSTGGGVGGGINFNPQGNNTNTNTGSPSPQQQQQH 746
Db	174 -----RNNS-----NQI 184
Qy	747 GITNKSFPLDAPRBDATLYLGDIQDHMTFQNLLAYEKIFPSR 791
Db	185 SLSN-SYQSTSOPRGDIALYLGDIQDHMTFQNLLAYEKIFPSR 228
XX	
RESULT 9	
ABP2160	
ID	ABP32160 standard, protein: 154 AA.
XX	
XX	Human glycoprotein-like ORP1133 protein, SEQ ID NO:2266.
XX	
KW	ORP; open reading frame; ORPX; drug screening; diagnosis;
KW	disease monitoring; cytokine; cell proliferation; cell differentiation;
KW	immune modulation; haemoptysis; regmatocyt; tissue growth;
KW	angiogenesis; activin; inhibin; chemotactic; hematologic; haemostatic;
XX	
XX	(first entry)

KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
 KW behavior; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnerable;
 KW vasoconstrictor; antidiabetic; cytostatic; nootropic;
 KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
 KW dermatologist; hypotensive; antithyroid; antiinflammatory; immunomodulator;
 KW cardiotonic; analgesic; virucide; antibacterial; fungicide.
 XX OS Homo sapiens.
 XX EP WO200190366-A2.
 XX ED 28-NOV-2001.
 XX BR 24-MAY-2001; 2001WO-US017076.
 XX PF 24-MAY-2000; 2000USP-0206690P.
 XX ID (CURA-) CURAGEN CORP.
 XX PA (CURA-) CURAGEN CORP.
 XX PI Leach MD, Shinkets RA.
 XX DR WPI; 2002-106200/14.
 XX DR N-PSDB; ABN76186.
 FP Novel human polypeptides and polynucleotides useful for diagnosing,
 FP preventing and treating cardiovascular diseases, neurodegenerative,
 FP hyperproliferative disorders and disorders related to organ
 FP transplantation.
 XX ES Claim 10; Page 830; 2508pp; English.
 CC Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN7587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF534 (collectively
 CC referred to as ORF) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of modulators of ORFX expression or
 CC activity, and methods of screening for modulators of ORFX expression or
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haemopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 CC and antinefective activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins
 CC nucleic acids and antibodies may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases
 XX Sequence 154 AA;

Query Match	7.2%	Score 345.5;	DB 5;	Length 154;
Best Local Similarity	44.6%	Pred. No. 5e18;	Mismatches 39;	Indels 39;
Matches	79;	Conservative	20;	Gaps 6;
Qy	351 GSSMAALKYPTPKNLKTKTSLRPEFTHENSSSBEIYLTKQOPPYKTDQQLSLSSTS 410			
Db	1 GSS-ALKYPTE-----ISR--TELEKNCVNTMYTSENNES-VREDDKPD----- 42			
Qy	411 TSGSGSGCGKFGGARRISDGING-----SLPREFLPLFSEEEHTHADP 454			
Db	43 ----HPDVTGPKNGKEGRGNDTSYSRAYTQLQNTEQTQSRSRFFSEDETASD 97			
Qy	455 IPSLIVSPGQSYTRDLFRNGGETWMLDCTCPDSDERMKLAKAFGTHPLTADIMQETR 511			
Db	98 IPSLISBQQTYYELFKGCDPWWLDSCSCTPDDEMRCIATKFGTHPLTADIRNQETR 154			
<hr/>				
RESULT 10				
ID	ABP06753	standard; protein; 107 AA.		
XX	XX	XX	XX	XX
AC	ABP06753;			
DT	24-JUN-2002	(first entry)		
DE	Human ORFX protein sequence SEQ ID NO:13488.			
KW	Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis; KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage; KW degenerative disorder; osteoarthritis; neurodegenerative disorder; KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus; KW hypertension; hypothyroidism; cholesterol ester storage disease; KW immune deficiency; immune disorder; infectious disease;			
KW	autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis; KW myasthenia gravis.			
OS	Homo sapiens.			
PN	WO200192521-A2.			
XX	XX	XX	XX	XX
PD	06-DEC-2001.			
PP	29-MAY-2001; 2001WO-US010836.			
PR	30-MAY-2000; 2000US-0206112P.			
PR	29-AUG-2000; 2000US-0228716P.			
PA	(CURA-) CURAGEN CORP.			
PI	Shinkets RA, Leach MD;			
DR	WPI; 2002-106308/14.			
DR	N-PSDB; ABN2505.			
XX	The present invention describes substantially purified human proteins CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1 CC in the specification). ABN15762 to ABN2732 encode the human ORFX CC proteins given in ABP0010 to ABP1500. ORFX proteins are useful for CC treating or preventing a pathology associated with an ORFX-associated CC disorder in humans, and in the manufacture of a medicament for treating a CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide CC sequences can be used in gene therapy. ORFX sequences can be used in the CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,			
PS	Disclosure; SEQ ID NO 13488; 1037pp; English.			
CC	Novel human polypeptides and polynucleotides useful for diagnosing, PT preventing and treating cardiovascular disease, neurodegenerative, PT hyperproliferative disorders and autoimmune disorders.			
CC	Novel human polypeptides and polynucleotides useful for diagnosing, PT preventing and treating cardiovascular disease, neurodegenerative, PT hyperproliferative disorders and autoimmune disorders.			
CC	Source of primers and probes, in the detection of ORFX genomic sequences CC or transcripts, in the identification and cloning of homologous CC sequences, in genetic diagnosis, and in forensic biology. The ORFX CC nucleic acids may additionally be used to produce transgenic animals CC which may be useful for studying the function and/or activity of ORFX CC protein, and in drug screening. The ORFX proteins may also be used as CC immunogens to generate specific antibodies, which are useful in the CC diagnosis, treatment and monitoring of ORFX-associated diseases			
CC	Sequence 154 AA;			

psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage, osteoarthritis, neurodegenerative disorders, disorders related to organ transplantation, cardiovascular diseases, diabetes mellitus, systemic lupus erythematosus, hypertension, hypothyroidism, cholesterol ester storage disease, various immune deficiencies and disorders, infectious diseases, autoimmune disorders such as multiple sclerosis, rheumatoid arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. ORPX proteins are also useful for treating burns, incisions, ulcers, for treating osteoporosis, bone degenerative disorders, or periodontal disease, and for gut protection or regeneration and treatment of lung or liver fibrosis, repurposing injury in various tissues and conditions resulting from systemic cytokine damage. N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at <http://wipo.int/pct/publication/pct-sequences>

XX
SQ Sequence 107 AA:
 Query Match 6.6%; Score 319.5; DB 5; Length 107;
 Best Local Similarity 62.9%; Pred. No. 2.8e-16;
 Matches 56; Conservative 21; Mismatches 11; Indels 1; Gaps 1;

Qy 790 SRSHTNYLAQLOVESFSNNKTEPFSKTTLIGTMLVPLNIVTGFLGMYRVPGEGGTNL 849
Ds 1 SRSHTNYLAQLOVESFSNNKTYEMLVGLKTMVIGTMLVPLNIVTGFLGMYRVPGE-NSSI 59
Qy 850 GMPFGTVGVLIPTIGTSPIAQWNIKKLN 878
Ds 60 AWWFGILGVLLLAVALGMFLASWIKRID 88

RESULT 11
ID ARN92441 standard; protein; 326 AA.
XX
AC ARN92441;
XX
DT 02-JUN-2005 (first entry)
DB M. xanthus protein sequence, seq id 11640.
XX
FW Transgenic plant; DNA replication; gene regulation; gene expression.
XX
OS Myxococcus xanthus.
PV US6833447-B1.
XX
PD 21-DEC-2004.
XX
PP 10-JUL-2001; 2001US-00902540.
XX
PR 10-JUL-2000; 2000US-0217881P.
XX
PA (MONSANTO TECHNOLOGY LLC).
XX
P1 Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;
XX
DR WPI; 2005-028716/03.
XX
PR New substantially purified Myxococcus xanthus nucleic acid molecule encoding a nitrite reductase, useful for determining gene expression, identifying mutations in a gene of interest, and for constructing mutations in a gene of interest.
XX
PS Example 2; SEQ ID NO 11640; 25pp; English.
XX
CC The invention relates to a substantially purified nucleic acid molecule encoding a nitrite reductase of SEQ ID NO. 11626. Further disclosed is a recombinant DNA construct for expression of a nitrite reductase gene in a plant cell, and a plant cell comprising the recombinant DNA construct.
XX

DE	Lactococcus lactis protein yfgQ.	664
XX	Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.	
KW		
XX		
XX	Lactococcus lactis; IL1403.	
OS		
XX		
PN	FR20170446-A1.	
XX		
XX	12-OCT-2001.	
PD		
XX		
PP	11-APR-2000; 20000FR-00004630.	
XX		
PR	11-APR-2000; 20000FR-00004630.	
XX		
PA	(INRG) INRA INST NAT RECH AGRONOMIQUE.	
XX		
PT	Bolotine A, Sorokine A, Renault P, Ehrlich SD;	
PI		
XX		
DR	WPI; 2002-043418/06.	
XX		
CC	New nucleotide sequence useful in the identification or Lactococcus lactis and related species.	
PS	Claim 6; SEQ ID NO 1665; 2504app; French.	
XX		
CC	The present invention is related to a Lactococcus lactis nucleotide sequence (ABA90521) and related proteins (ABB51301-AB855621). The nucleic acid sequence is useful in the detection and/or amplification of nucleic acid sequence, particularly to identify Lactococcus lactis or related species. The proteins of the invention are useful for the biosynthesis or biodegradation of a composition of interest. The invention helps research in lactic bacteria, particularly useful in the production of yogurt and cheese. Note: The sequence data for this patent is based on equivalent patent WO200177334 (published 18-OCT-2001) which is available in electronic format directly from WIPO at http://wipo.int/pub/published_pct_sequences . (Updated on 29-AUG-2003 to standardise OS field)	
CC		
XX		
SQ	Sequence 1072 AA:	
Query Match	3.9%: Score 186.5; DB 5; Length 1072;	
Best Local Similarity	20.8%; Pred. No. 0.00012;	
Matches	191; Conservative 136; Mismatches 363; Indels 235; Gaps 45;	
Qy	1 MDSSEYYQ-NSTNQPIRSDEVLDIHRNQITN-DCAIISDSEDEBLSKSELESEVYKSE 58	
Db	305 VASSEQTQANSASLYPISEASVTDNTLANSISSDSSSISOTP-----NSQ 352	
Qy	59 KQQHQHQEQTSDNAPLTKSSSIIKKSNLNDTIDTNPMSLGGDDTINSKRRN 118	
Db	353 SGASSTAETSDENSENSNLSSQNQNSNSN-SKEKD--SNGQSLG---SSMSNEEBSHN 406	
Qy	119 MSSLRKDFKLKDNTNNTNTHLAIPITPIPAPITNNANKSRRKSQDENLPPLIKKK 178	
Db	407 SSNI-----NFTNNSSET-----TNIILPENPENPES-----VSDQ 437	
Qy	179 TIGRNNNFENDLVSPMTKRNEDTNTS-----TPANHAKIGIG----- 222	
Db	438 TSEBAAATNSNSTSISPNNTSSEASATISDSENVAYANNLSAVNNSSSVLSST 497	
Qy	223 --ATTLGV---GTCTTATAATAAGGRPSRSSDSEADASHASRSQETEDVCPMVG 276	
Db	498 STADNGIQSSESDNTKDSSEISSTSGAFLSSQNSPSSASTNSNISIISLSPNTSISSTVL 557	
Qy	277 DHIRVNGIDPDBIDEFIR----PEREAVLQKOMIAKNTLDEP--ONLSKNN---T 325	
Db	558 ES-TTSSNSFSNVAEVANNLSATNNSSSSLSSTSTADN-LENNOQGSNLDKSEISI 615	
Qy	326 TSSASRHPYTHHSNNNNKANGGCGSSMAALKYTPKNLAKTUSREPFETHNSSEEEI 385	

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2006, 13:59:19 : Search time 314 Seconds
(without alignments)

2716.129 Million cell updates/sec

Title: US-10-018-10SA-4

Perfect score: 4820

Sequence: 1 MSDSEYYQNTTNTQPIPRS.....LGUKKHHGGNKSITISPPNKYE 922

Scoring table: BLOSUM62
Gapext 10.0 , Gapext 0.5

Searched: 2849598 seqs, 92501592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database : Uniprot_7.2.*

1: uniprot_sprot:
2: uniprot_trembl:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	4723	98.0	932	2	Q5AJ5_CANAL	05aj5_canal
2	4712.5	97.8	929	2	Q5AJ3_CANAL	05aj3_candida alb
3	2056	42.7	936	2	Q5BL8_DEBHA	05bl8_debaromyces
4	1683.5	34.9	903	2	Q5V15_CANGA	008269_saccharomyces
5	1680	34.9	859	1	Q5R11_YEAST	043553_saccharomyces
6	1666	34.6	858	1	Q5R22_YEAST	075au1_asbya_goss
7	1641.5	34.1	799	2	Q5CP3_KLUHG	06cp63_kluveromyces
8	1603	33.3	778	2	Q5C63_KLUHG	06cp66_yarrowia_li
9	1262.5	26.2	605	2	Q5C46_YARL1	013657_chirossach
10	1173	24.3	803	2	Q5C67_SCHPO	013658_chirossach
11	1141.5	23.7	699	2	Q5D40_SCHPO	041w68_schizosaccharomyces
12	1059	22.8	663	2	Q5W15_ASPPU	049j10_yarrowia_ma
13	1044	21.7	460	2	Q5C70_YARL1	04phq2_ustilago_ma
14	1002	20.8	881	2	Q5PH2_JUSTRA	02u117_aspergillus
15	990	20.5	756	2	Q5UTL7_ASPOR	055p24_cryptococcus
16	975.5	20.2	791	2	Q5KE2_CRYNE	05ked2_cryptococcus
17	975.5	20.2	791	2	Q5KE2_CRYNE	05ked2_cryptococcus
18	971.5	20.2	617	1	Q5E16_SCIB20	041i33_schizosaccharomyces
19	935	19.4	629	2	Q5J03_GIB22	041i33_gibberella
20	916	19.0	836	2	Q5J07_GIB22	041q07_neurospora
21	905.5	18.8	822	2	Q7SEN2_NEUCR	07sen2_neurospora
22	900.5	18.7	877	2	Q4X211_ASPPU	06c1211_aspergillus
23	888.5	18.4	962	2	Q6CTG3_KLUHG	06c193_kluveromyces
24	887.5	18.4	884	2	Q2T2D7_ASPOR	02t2d7_aspergillus
25	885	17.9	969	1	Q5R22_YEAST	05f579_saccharomyces
26	836	17.8	797	2	Q5903_CANAL	05903_candida_alb
27	842.5	17.5	1107	2	Q5F15_CANGA	06f1ms_candida_glia
28	835	17.3	836	2	Q5BD4_EMENI	05b0s4_aspergillus
29	823.5	17.1	1000	2	Q6BP3_DEBRA	06bp03_debaromyces
30	784.5	16.3	1001	2	Q75DD5_ASHGO	075dd5_asbya_goss
31	670.5	13.9	399	2	Q4W936_ASPPU	04w936_aspergillus
32	621.5	12.9	414	2	Q5AK0_EMENI	05axp0_aspergillus
33	618.5	12.8	409	2	Q2U08_ASPOR	02u108_aspergillus
34	615.5	12.8	366	2	Q5BB9_EMENI	05bb9_aspergillus
35	564	11.3	291	2	Q2U20_ASPOR	02u20_aspergillus
36	363.5	7.5	990	2	Q5E15_CRYNE	0561j5_cryptococcus
37	362	7.5	989	2	Q5KQ98_CRYNE	05kq98_cryptococcus

38 337.5 7.0 387 2 Q54YH3_DICDI 054yw3 dictyostelia
39 327 6.8 625 2 Q54BR2_DICDI 054er2 dictyostelia
40 313 6.5 1423 2 Q5AP177_USTNA 04p177 ustilago ma
41 298.5 6.2 427 2 Q51BW1_ENTHI 051bw1 entamoeba_h
42 298 6.2 387 2 Q554BL_SYNTH3 0554bl synchrocyt
43 266.5 5.5 354 2 Q4APB3_9CHUB 04apb3 chlorobium
44 263.5 5.5 374 2 Q43HE4_9CHUB 043he4 chlorobium
45 259 5.4 385 2 Q7NIU6_GLOVT 07niu6 gloeobacter

ALIGNMENTS

RESULT 1									
ID	Q5AJ5_CANAL	PRELIMINARY;	PRP;	932 AA.	AC	Q5AJ5_CANAL	PRELIMINARY;	PRP;	932 AA.
DT	26-APR-2005; integrated into UniProtKB/TREMBL.				DT	26-APR-2005; sequence version 1.			
DT	07-FEB-2006; entry version 6.				DE	Hypothetical protein ARL1.			
GN	Name=ARL1; ORPName=Cao19_9175;				IS	Candida albicans SC314.			
OC	OC				OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; mitosporic Saccharomycetales; Candida.			
OX	NCBI_TaxID:237561;				OX				
RN	[1]				RN	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].			
RC	STRAINS:SC314.				RC				
RA	Magee B.B., Newport G., Thorssen Y.R., Chibana H., Dungan J., Kalman S., Magee P.T., Davis R.W., Scheuer S.				RA	The diploid genome sequence of <i>Candida albicans</i> . Proc. Natl. Acad. Sci. U.S.A. 101:7129-7134 (2004).			
RL	!- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.				RL				
CC	Copyright by the Uniprot Consortium, see http://www.uniprot.org/terms Distributed under the Creative Commons Attribution-NonDerivs License				CC				
CC	DR	EMBL: AAC001000014; EAU02768.1: -; Genomic DNA.			CC	DR	GO; GO:0016020; C-membrane; IEA.		
CC	GO	GO:0016020; C-membrane; IEA.			CC	GO	GO:0046872; P-metal ion transporter; activity; IEA.		
CC	DR	DR			CC	DR	InterPro:IPR02523; Mg2+-transport; CoRA.		
CC	PFam: PF01544; CorA_1.				CC	KW	Hypothetical protein.		
CC	SO	SEQUENCE 932 AA: 101725 MNW; 5BDB2A12BA00B17 CRC64;			CC	Query Match	98.0%; Score 4723; DB 2; Length 932; Matches 916; Conservative 916; P_mismatches 2; P_alignments 2; Gaps 6;		
CC	1	MSDSEYYQNTTNTQPIPRS.....LGUKKHHGGNKSITISPPNKYE 922			QY	1	MSDSEYYQNTTNTQPIPRS.....LGUKKHHGGNKSITISPPNKYE 922		
CC	1	MSDSEYYQNTTNTQPIPRS.....LGUKKHHGGNKSITISPPNKYE 922			Db	1	MSDSEYYQNTTNTQPIPRS.....LGUKKHHGGNKSITISPPNKYE 922		
CC	60 ---QQQHQHQTSDNAAPLTKRGSSTIKKSNLTDRTNTPMSLGGDDTINSCKARN 116				QY	60 ---QQQHQHQTSDNAAPLTKRGSSTIKKSNLTDRTNTPMSLGGDDTINSCKARN 116			
CC	61 QQQQQHQHQTSDNAAPLTKRGSSTIKKSNLTDRTNTPMSLGGDDTINSCKARN 120				Db	61 QQQQQHQHQTSDNAAPLTKRGSSTIKKSNLTDRTNTPMSLGGDDTINSCKARN 120			
CC	117 YMSSLRKDFYLKDNTDNSTNNNTHLAIP1PIPPIPITANSRSKOLENLPLIK 176				QY	117 YMSSLRKDFYLKDNTDNSTNNNTHLAIP1PIPPIPITANSRSKOLENLPLIK 176			
CC	121 YMSSLRKDFYLKDNTDNSTNNNTHLAIP1PIPPIPITANSRSKOLENLPLIK 180				Db	121 YMSSLRKDFYLKDNTDNSTNNNTHLAIP1PIPPIPITANSRSKOLENLPLIK 180			
CC	177 KTKTRGRNHNPFENDLVSPATKTKNDSEPDITSTPAHMLGATTLGVGT---TA 234				QY	177 KTKTRGRNHNPFENDLVSPATKTKNDSEPDITSTPAHMLGATTLGVGT---TA 234			
CC	181 KTKTRGRNHNPFENDLVSPATKTKNDSEPDITSTPAHMLGATTLGVGT---TA 240				Db	181 KTKTRGRNHNPFENDLVSPATKTKNDSEPDITSTPAHMLGATTLGVGT---TA 240			
CC	215 TATAAAAGRGPSSSIDSEA DSASHSRSSQETEDVCPMVGDFHTRVNGIDFDEDETR 294				QY	215 TATAAAAGRGPSSSIDSEA DSASHSRSSQETEDVCPMVGDFHTRVNGIDFDEDETR 294			
CC	241 TATATAGRGPSSSIDSEA DSADHARSQETEDVCPMVGDFHTRVNGIDFDEDETR 300				Db	241 TATATAGRGPSSSIDSEA DSADHARSQETEDVCPMVGDFHTRVNGIDFDEDETR 300			
CC	255 BERBAYLQKOMIAKNTLRIDEFQNLSSKNTSGASRPHYDHSHNNKKNGDG-SGS 353				QY	255 BERBAYLQKOMIAKNTLRIDEFQNLSSKNTSGASRPHYDHSHNNKKNGDG-SGS 353			
CC	301 EEREAVLQKOMIAKNTSIRDEFQNLSSKNTSGASRPHYHHSHNNKKNGDGSGSS 360				Db	301 EEREAVLQKOMIAKNTSIRDEFQNLSSKNTSGASRPHYHHSHNNKKNGDGSGSS 360			
CC	37 362 7.5				CC	37 362 7.5			

Qy	354	MALKYTPKNIKTKTSLSERPEHTENSSSERIVELTKKOPPKYIDQQLSLSSTT-	411	Query Match	97.8%	Score 4712.5;	DB 2;	Length 929;
Db	361	MAALKYTPKNIKTKTSLSERPEHTENSSSERIVELTKKOPPKYIDQQLSLSSTTSG	420	Best Local Similarity	98.3%;	Pred. No. 9.4e-260;		
		Matches 914; Conservative 0; Mismatches 27; Indels 9; Gaps 6;						
Qy	412	SGSGSCQVKKFGEARISDGINGGSLPDRFLSHSESEETIADPIPILSPGOSVRLDPN	471	Qy	1	MSDESEYYQNNTMNPQIPRSEDEVLDLDRHQNTDITCDAISDSEDELKSELESEVTSER-	59	RESULT 3
Db	421	SGSGSCQVKKFGEARISDGINGGSLPDRFLSHSESEETIADPIPILSPGOSVRLDPN	480	Db	1	MSDESEYYQNNTMNPQIPRSEDEVLDLDRHQNTDITCDAISDSEDELKSELESEVTSER-	60	Q6BL48_
Qy	472	GERTWILDCTCPTOSEKMLAKAGIPLAEDTRQETREKELKSIFCPIPTREAD	531	Qy	60	COOHQQETSDNAPLTRKSSSERIKKSNTUDRITNMSLSCGDTTNSCHKARYNM	119	ID
Db	481	GERTWILDCTCPTDSEMKMLAKAGIPLAEDTRQETREKELKSIFCPIPTREAD	540	Db	61	OQHHQETSDNAPLTRKSSSERIKKSNTUDRITNMSLSCGDTTNSCHKARYNM	120	AC
Qy	532	KESEDYLEFINVIVPFHDGLTHPSPISHANVRVRLQDIDYDSADMCLVLTIDE	591	Qy	120	SSLRDKDYLKONTDDNSTNNTHLAIPPIPIPTITANTSRSRQESOLENLPLIKKTT	179	AC
Db	541	KESEDYLEFINVIVPFHDGLTHPSPISHANVRVRLQDIDYDSADMCLVLTIDE	600	Db	121	SSLRDKDYLKONTDDNSTNNTHLAIPPIPIPTITANTSRSRQESOLENLPLIKKTT	180	DT
Qy	592	ITDGPAPVTHGIEYADAEVIAVTAUDTESSLORIGESRKVATLRLISGRADVTK	651	Qy	180	IGRANSNNPFDNDLVSPTKAKTMSDITIN-TOTTANHKLGLGATTLGVTNTATA	238	16-AUG-2004;
Db	601	ITDGPAPVTHGIEYADAEVIAVTAUDTESSLORIGESRKVATLRLISGRADVTK	660	Db	181	IGRANSNNPFDNDLVSPTKAKTMSDITIN-TOTTANHKLGLGATTLGVTNTATA	240	integrated into UniProtKB/Trembl.
Qy	652	MFAKRCQEANSSGGYCYORINLQOQQQQ--APPENPNITSPINSTLNLSLGTSRG	709	Qy	239	TAAA -- -CRPRSSSIDSEADHASRSQETEEDVCFPMVGDHIRVNGIDPDEIDEPIRE	295	DT 16-AUG-2004,
Db	661	MFAKRCQEANSSGGYCYORINLQOQQQQ-APPENPNITSPINSTLNLSLGTSRG	720	Db	241	TATAGGCCRSPRSSIDSEADHASRSQETEEDVCFPMVGDHIRVNGIDPDEIDEPIRE	300	sequence version 1.
Qy	710	GVGVGTGNGRPNTGNNTNTNTGSPSPQQQQGTTKNSPPDPAPRADALYG	769	Qy	296	BREAYLOXQKQIARNIRIDPONLSKNTTSAASHPYTHRSNNNNKRNGDGD-GGSNM	354	DB 16-AUG-2004,
Db	721	GVGVGTGNGRPNTGNNTNTNTGSPSP-OOQGTTKNSPPDPAPRADALYG	779	Db	301	EREAYLOXQKQIARNIRIDPONLSKNTTSAASHPYTHRSNNNNKRNGDGSNSSM	360	OrderedlocusNames=DEHAOF1776g;
Qy	770	DIQDHLMITPQNLAYEKIFSRSHSNYLAQOVSNSNNKITEMPSKTLIGTMVLPI	829	Qy	355	AALKTPKNLKKLTSRPEPHENISSSERIYELKTKQKPPYDOLSLTSSTSSTSGS	414	
Db	780	DIQDHLMITPQNLAYEKIFSRSHSNYLAQOVSNSNNKITEMPSKTLIGTMVLPI	839	Db	361	AALKTPKNLKKLTSRPEPHENISSSERIYELKTKQKPPYDOLSLTSSTSSTSGS	420	
Qy	830	LVTGLFGKVNRYPGEGGTNLGMPFGIVYUFLITIGSFIPAQWMLKLANSIEGQNGNTR	889	Qy	415	GSGQVKFKEGARIISDGINGSISPDPSLSPHSBSSETHADPISLVSPGOSVRLPREGES	474	
Db	840	LVTGLFGKVNRYPGEGGNLGMFGIVYUFLITIGSFIPAQWMLKLANSIEGQNGNTR	899	Db	421	GSGQVKFKEGARIISDGINGSISPDPSLSPHSBSSETHADPISLVSPGOSVRLPREGES	480	
Qy	890	P1PNMSSRSRSRSGLKHGKNSKTSISPKRYE	922	Qy	475	TWLLDCTCPDSEKMKLAKAIPGTHPLTAEDIRHOETREKVELFKSYKCPHTEADEKES	534	
Db	900	P1PNMSSRSRSRSGLKHGKNSKTSISPKRYE	932	Db	481	TWLLDCTCPDSEKMKLAKAIPGTHPLTAEDIRHOETREKVELFKSYKCPHTEADEKES	540	
Qy				Qy	535	EDYLEPINYIVVPHDGLLTHPSPISHANVRVRLQDIDYDSADMLCYALDIDTD	594	
Db				Db	541	EDYLEPINYIVVPHDGLLTHPSPISHANVRVRLQDIDYDSADMLCYALDIDTD	600	
Qy				Qy	595	GPAPIHGLEYADAEVDAVFTARDTDFDSMORIGERSRKYNTMLRJLSKADYVTKMPA	654	
Db				Db	601	GPAPIHGLEYADAEVDAVFTARDTDFDSMORIGERSRKYNTMLRJLSKADYVTKMPA	660	
Qy				Qy	655	KRCQEBAANSGGYCYORINLQOQQQQ--APPNNPPLTSPINSTLNLSGTTGCGV	712	
Db				Db	661	KRCQEBAANSGGYCYORINLQOQQQQ-QAPPNNPPLTSPINSTLNLSGTTGCGV	720	
Qy				Qy	713	VGGINGNPQGANTNTNTGSESPQQQQGTTGTTGTTGTTGTTGTTGTTGTTG	772	
Db				Db	721	VGGINGNPQGANTNTNTGSESPQQQQGTTGTTGTTGTTGTTGTTGTTGTTG	779	
RN				Qy	773	DHITMEPQNLAYEKIFSRSHSNYLAQOVSNSNNKITEMPSKTLIGTMVLPLNYT	832	
NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA)				Db	780	DHITMEPQNLAYEKIFSRSHSNYLAQOVSNSNNKITEMPSKTLIGTMVLPLNYT	839	
STRAIN=SC5114;				Qy	781	GLFGNTRVPBEGGTNLGMPFGIVYUFLITIGSFIPAQWMLKLANSIEGQNGTR	892	
PUBMED=15123810;				Db	840	GLFGNTRVPBEGGTNLGMPFGIVYUFLITIGSFIPAQWMLKLANSIEGQNGTR	899	
HYPOTHETICAL PROTEIN ARL1.				Qy	893	NHSRSPRSRSGLKKHGCGNTSISPKRYE	922	
NAME=ARL1; ORTNAME=Ca019_1607;				Db	900	NHSRSPRSRSGLKKHGCGNTSISPKRYE	939	
CANDIDA ALBICANS SC5114.								
Eukaryota; Fungi; Ascomycota; Saccharomyctina; Saccharomycetes;								
SACCHAROMYCETES; Candida; Candida; Candida.								
NCBI TAXID=237561;								
[1]								
RP								
NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).								
STRAIN=SC5114;								
PUBMED=101073; ONAS_04016481_01;								
JONES T., FEDERSTEIN N.A., CHIBABA H., DUNGAN J., KALMAN S., NAGEE P.T.,								
DAVIS R.W., NEWPORT G., MORSTENSON Y.R., AGABIAN N., NAGEE P.T.,								
RAT								
THE DIPLOID GENOME SEQUENCE OF CANDIDA ALBICANS. *								
PROC. NATL. ACAD. SCI. U.S.A. 101:7329-7334 (2004).								
- - CAUTION: The sequence shown here is derived from an								
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is								
PRELIMINARY DATA.								
CC								
CC		Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms						
CC		DISTRIBUTED UNDER THE CREATIVE COMMONS ATTRIBUTION-NO DERIVS LICENSE						
CC		EBML: AAC001000013; EMBL: 028956; 1- Genomic DNA.						
DR		GO: GO:0016020; C: membrane; IAA.						
DR		GO: GO:046873; P: metal ion transporter activity; IEA.						
DR		GO: GO:030001; P: metal ion transport; IEA.						
DR		InterPro: IPR00253; Mg+_transporter.						
DR		Protein: P01344; CorA; 1.						
RW		Hypothetical protein.						
SQ		SEQUENCE 929 AA: 103372 MW: 98ADF46242559SE4 CRC64:						

OS	Debaromyces hansenii (Yeast) (Torulaspora hansenii).	Db	453 VPKGSQSVSDLFPRNGEGTWLDCVCPDPAEMKVKAKPGHPLTAEDRQETREKTELFR	512
OC	Bukaryote; Fungi; Ascomycota; Saccharomycoctina; Saccharomycetes.	Qy	519 SYTFYCFTTFADEKESEDYLPIPNVYTIVFHGDILTHFPISHANVRRTYRQLDDYD	578
OC	Saccharomyctaceae; Saccharomyces.	Qy	513 NYFYFCFTTFADEKESEDYLPIPNVYTIVFHGDILTHFPISHANVRRTYRQLDDYD	572
NCBI_TaxID=1955;	[1]	Db	513 NYFYFCFTTFADEKESEDYLPIPNVYTIVFHGDILTHFPISHANVRRTYRQLDDYD	572
RN	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].	Qy	519 VSADMCYCALDEDTDSFAPTHSIEYEADAIIDANTPARDTDFSSMLORGESBRKVT	638
RC	STRAIN=ATCC 36239 / CBS 767/	Db	513 VSADMCYCALDEDTDSFAPTHSIEYEADAIIDANTPARDTDFSSMLORGESBRKVT	632
RR	Pubmed=15245929; DOI=10.1038/nature02579;	Qy	513 VSADMCYCALDEDTDSFAPTHSIEYEADAIIDANTPARDTDFSSMLORGESBRKVT	632
RA	Dufont B., Sherman D., Fischer G., Durrens P., Casaregoia S., Talia E., Lafontaine D., Montjoly J., March C., Neuveglise C., Talia E., Goffard N., De Moul L., Aigle M., Anthoud V., Babour A., Barbe V., Barnay S., Blanchard S., Beckerich J.-M., Beyne E., Bleckmann C., Boisrame A., Boyer J., Cattolico L., Confaquieri F., De Bakriavar A., Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A., Hantraye P., Hennequin C., Junnaius N., Jouet P., Kachouri R., Kerrest A., Koszuil R., Lemire M., Lesur I., Maler H., Niclau J.-M., Nikolski M., Oztas S., Oztas S., Ozler S., Oztas S., Peillen S., Potier S., Richard G.-P., Straub M.-L., Suleau A., Swennen D., Tekaya F., Wessolowski M., Westhof E., Wirth B., Zaniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A., Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissbach J., Wincker P., Souciet J.-L., "Genome evolution in yeasts." Nature 430:35-44 (2004).	Db	513 VSADMCYCALDEDTDSFAPTHSIEYEADAIIDANTPARDTDFSSMLORGESBRKVT	632
RA	Peillen S., Potier S., Richard G.-P., Straub M.-L., Suleau A., Swennen D., Tekaya F., Wessolowski M., Westhof E., Wirth B., Zaniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A., Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissbach J., Wincker P., Souciet J.-L., "Genome evolution in yeasts." Nature 430:35-44 (2004).	Qy	513 VSADMCYCALDEDTDSFAPTHSIEYEADAIIDANTPARDTDFSSMLORGESBRKVT	632
RA	Copyright by the UniProt Consortium, see http://www.uniprot.org/terms Distributed under the Creative Commons Attribution-NoDerivs license	Db	513 VSADMCYCALDEDTDSFAPTHSIEYEADAIIDANTPARDTDFSSMLORGESBRKVT	632
CC	EBML; CR182138; CAG89451_1; -; Genomic_DNA.	Qy	639 LMRLLGDAYTKTSPFKRKQCEA-----NSSTGNS---GQOQAPP	683
CC	GO: GO:0016020; C:membrane; IEA.	Db	633 LMRLLGDAYTKTSPFKRKQCEAANNQINIFQAHFNNGNIGNTY-SQTSGNSASQSHHSGS	692
DR	GO: GO:0046873; P:metal ion transporter activity; IEA.	Qy	684 PPPNPITSPINSTNLNSLGTSGCGVGCGNFGNPICGNTNTTTSPPSQOQ	743
DR	GO: GO:0030001; P:metal ion transporter activity; IEA.	Db	693 YTHGTYNNPSPPQGANNS-----NPQGNNSNPQYANNNSPQYSNNPAPQY	743
DR	InterPro:IPR02523; Mg2+-transportCora.	Qy	744 QQ-----HGTTKNSP-----PIPD-APIPDAIYLGDIQDHITMPNLLA	784
DR	PFam: PF01544; Cora_1.	Db	744 NGGGPGVTFGPGCVSNMMTIDRGAFYMKQTQDPTRADALYLGDIQDHVMTFQNLVA	803
DR	Complete proteome; SEQUENCE 936 AA;	Qy	785 TEKTSRSHSTYLQLOQVESENNSNSINKTENPSKTLISTIMLVLPSNLYTGLGRMVNPSC	844
DR	SEQUENCE 936 AA;	Db	804 YEKTSRSHSYALQLOQVESENNSYKTYALSNVMTGIFTFLPFLVNTGIFGRNVVPQQ	863
CC	Copyright by the UniProt Consortium, see http://www.uniprot.org/terms Distributed under the Creative Commons Attribution-NoDerivs license	Qy	845 GETN-LGWFFGIVLULIGSPTIAQWLKLNSIEGONGNRPIFHNSRSIRS	902
CC	CC	Db	864 GDDHFGLSHFFGIGVCAUTIVAVSARTMILRVERTYDEEN-----TRKSXKS	914
DR	MSDSEYNTNTNQPIPR---SDEVLDIHDNQINTN-----DCAITSDEDEL 45	Qy	903 LGLKXH--GGNKSLISPNKVE	922
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	915 YELKTRTRKEAAKTSISPNKYD	936
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	RESULT 4	
DR	MSDSEYNTNTNQPIPR---SDEVLDIHDNQINTN-----DCAITSDEDEL 45	Db	Q6FV15_CANGA	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	ID Q6FV15_CANGA	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	PRELIMINARY;	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	PRP; 803 AA.	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	845 GETN-LGWFFGIVLULIGSPTIAQWLKLNSIEGONGNRPIFHNSRSIRS	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	864 GDDHFGLSHFFGIGVCAUTIVAVSARTMILRVERTYDEEN-----TRKSXKS	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	903 LGLKXH--GGNKSLISPNKVE	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	915 YELKTRTRKEAAKTSISPNKYD	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	915 YELKTRTRKEAAKTSISPNKYD	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	RESULT 4	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	Q6FV15_CANGA	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	ID Q6FV15_CANGA	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	PRELIMINARY;	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	PRP; 803 AA.	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	845 GETN-LGWFFGIVLULIGSPTIAQWLKLNSIEGONGNRPIFHNSRSIRS	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	864 GDDHFGLSHFFGIGVCAUTIVAVSARTMILRVERTYDEEN-----TRKSXKS	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	903 LGLKXH--GGNKSLISPNKVE	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	915 YELKTRTRKEAAKTSISPNKYD	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	915 YELKTRTRKEAAKTSISPNKYD	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	RESULT 4	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	Q6FV15_CANGA	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	ID Q6FV15_CANGA	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	PRELIMINARY;	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	PRP; 803 AA.	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	845 GETN-LGWFFGIVLULIGSPTIAQWLKLNSIEGONGNRPIFHNSRSIRS	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	864 GDDHFGLSHFFGIGVCAUTIVAVSARTMILRVERTYDEEN-----TRKSXKS	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	903 LGLKXH--GGNKSLISPNKVE	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	915 YELKTRTRKEAAKTSISPNKYD	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	915 YELKTRTRKEAAKTSISPNKYD	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	RESULT 4	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	Q6FV15_CANGA	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	ID Q6FV15_CANGA	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	PRELIMINARY;	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	PRP; 803 AA.	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	845 GETN-LGWFFGIVLULIGSPTIAQWLKLNSIEGONGNRPIFHNSRSIRS	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	864 GDDHFGLSHFFGIGVCAUTIVAVSARTMILRVERTYDEEN-----TRKSXKS	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	903 LGLKXH--GGNKSLISPNKVE	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	915 YELKTRTRKEAAKTSISPNKYD	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	915 YELKTRTRKEAAKTSISPNKYD	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	RESULT 4	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	Q6FV15_CANGA	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	ID Q6FV15_CANGA	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	PRELIMINARY;	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	PRP; 803 AA.	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	845 GETN-LGWFFGIVLULIGSPTIAQWLKLNSIEGONGNRPIFHNSRSIRS	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	864 GDDHFGLSHFFGIGVCAUTIVAVSARTMILRVERTYDEEN-----TRKSXKS	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	903 LGLKXH--GGNKSLISPNKVE	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	915 YELKTRTRKEAAKTSISPNKYD	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	915 YELKTRTRKEAAKTSISPNKYD	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	RESULT 4	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	Q6FV15_CANGA	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	ID Q6FV15_CANGA	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	PRELIMINARY;	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	PRP; 803 AA.	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	845 GETN-LGWFFGIVLULIGSPTIAQWLKLNSIEGONGNRPIFHNSRSIRS	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	864 GDDHFGLSHFFGIGVCAUTIVAVSARTMILRVERTYDEEN-----TRKSXKS	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	903 LGLKXH--GGNKSLISPNKVE	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	915 YELKTRTRKEAAKTSISPNKYD	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	915 YELKTRTRKEAAKTSISPNKYD	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	RESULT 4	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	Q6FV15_CANGA	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	ID Q6FV15_CANGA	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	PRELIMINARY;	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	PRP; 803 AA.	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	845 GETN-LGWFFGIVLULIGSPTIAQWLKLNSIEGONGNRPIFHNSRSIRS	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	864 GDDHFGLSHFFGIGVCAUTIVAVSARTMILRVERTYDEEN-----TRKSXKS	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	903 LGLKXH--GGNKSLISPNKVE	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	915 YELKTRTRKEAAKTSISPNKYD	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	915 YELKTRTRKEAAKTSISPNKYD	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	RESULT 4	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	Q6FV15_CANGA	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	ID Q6FV15_CANGA	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	PRELIMINARY;	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	PRP; 803 AA.	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	845 GETN-LGWFFGIVLULIGSPTIAQWLKLNSIEGONGNRPIFHNSRSIRS	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	864 GDDHFGLSHFFGIGVCAUTIVAVSARTMILRVERTYDEEN-----TRKSXKS	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	903 LGLKXH--GGNKSLISPNKVE	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	915 YELKTRTRKEAAKTSISPNKYD	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	915 YELKTRTRKEAAKTSISPNKYD	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	RESULT 4	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	Q6FV15_CANGA	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	ID Q6FV15_CANGA	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	PRELIMINARY;	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	PRP; 803 AA.	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	845 GETN-LGWFFGIVLULIGSPTIAQWLKLNSIEGONGNRPIFHNSRSIRS	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	864 GDDHFGLSHFFGIGVCAUTIVAVSARTMILRVERTYDEEN-----TRKSXKS	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	903 LGLKXH--GGNKSLISPNKVE	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	915 YELKTRTRKEAAKTSISPNKYD	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	915 YELKTRTRKEAAKTSISPNKYD	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	RESULT 4	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	Q6FV15_CANGA	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	ID Q6FV15_CANGA	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	PRELIMINARY;	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	PRP; 803 AA.	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	845 GETN-LGWFFGIVLULIGSPTIAQWLKLNSIEGONGNRPIFHNSRSIRS	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	864 GDDHFGLSHFFGIGVCAUTIVAVSARTMILRVERTYDEEN-----TRKSXKS	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	903 LGLKXH--GGNKSLISPNKVE	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	915 YELKTRTRKEAAKTSISPNKYD	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	915 YELKTRTRKEAAKTSISPNKYD	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	RESULT 4	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	Q6FV15_CANGA	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	ID Q6FV15_CANGA	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	PRELIMINARY;	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	PRP; 803 AA.	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Db	845 GETN-LGWFFGIVLULIGSPTIAQWLKLNSIEGONGNRPIFHNSRSIRS	
DR	MSDNEPSVSESELETAAMNTGEVLDIHDNQIPLSPLIRGTCGGMGSANVDDDDDDY 60	Qy	864 GDDHFGLSHFFGIGVCAUTIVAVSARTMILRVERTYDEEN-----TRKSXKS	
DR				

RW	Complete Proteome	89694 MW;	E741B17C22E235A4 CRC64;		
50	SEQUENCE	803 AA;			
Query Match	Score 34.9%	DB 2;	Length 803;		
Best Local Similarity	43.0%;	Pred. No. 2.9-87;			
Matches 413; Conservative	92; Mismatches 215;	Indels 241;	Gaps 31;		
Qy	38 SDSEDE---DELSELESEVSEKQQ---OHHOBITSDAKPLTKSGSSIKKKSN 88				
Db	3 SSESSEPHSPQHSTVAESMKVSKNTDNELYLTHORHQI-----VPPRLGDSFGSSN 56				
Qy	89 LDTK-----DRYTNP---MSLGGDDTINSKHNRNTNMSSLRKFY-LKDNTDNSTNN 139				
Db	57 LNARHMLDGRANEDPGMANIEGADS-----DCPELGKDHKLRSALQNSMR 105				
Qy	140 RTHLALPI---PIPIPPLIT-NANKSERKSQLENPLTIKTKGRNFSANFENDLVP 195				
Db	106 PSRLAITYVHNVQVPAATITNSELKRNVAIAASP-----TKTRGQSNTSFLTDKTP 160				
Qy	196 MTKMKTNDSEDTNTSTTAHKRKGIGATTLVGTGTTATAAGGRPSRSSID---	252			
Db	161 L-KRTMSRINSTNSAVSSNRLNL-----AVNSISIHKSYDDN 199				
Qy	253 -----SEADSHASRSSEETEDVCFPAVDG-HIRYNGIDFDEDFIREEEKEY 301				
Db	200 LKKKTRTASSPDNSSQARESCTEEDVCFPMPPOLISRYNGIDPDELPFEEQVNE---	256			
Qy	302 LOKOMIAKHNLLIDEFONLSKNTTSGASRHPYHNNNNNTKNGGDGGSSM-----	354			
Db	257 -QKLLC-----EISLAAQKSG-----SNGTQAASSSSANSYRS 289				
Qy	355 -ALKYTPNKLKTYL--RFPFTHESSSEETYB-----LTKKQPPYKDDQ 401				
Db	290 EAALKYTPH--LRKOLSPKSELSNSNSNPEDNOPOGVSFGRNNTIEGEDEP---YEAR 345				
Qy	402 LSUTSSTSSTSSGSGOVKEGGARLSDGINGSPLDPHSESESETHTADYLWTP 461				
Db	346 PQAUTOPENPQYDGSKF-----APDRSFSSMAMESTHTADYLWTP 391				
Qy	452 GQVRDLPINGBTWMDCTPDTSEKMKLAKAFGIPHLTAEDIQETREKVELFKSY 521				
Db	392 GQFFPLPFGPPTWMDCTPDTSEKMKLAKAFGIPHLTAEDIQETREKVELFKSY 451				
Qy	522 KPCPTEPKADESKEDYLAPINVYIVWHDGLUTPHFEPISHRANVRURQRLDYDVS 581				
Db	452 FWCPTPTEPKADESKEDYLAPINVYIVWHDGLUTPHFEPISHRANVRURQRLDYDVS 511				
Qy	582 DMCYALIDETGDPATHGIGYEADAEVATAEDTDTFSMLORGESREKVKMLR 641				
Db	512 DMCYALIDDTDSAPYTSRSLTEADSTEDSPTMARDLDFAMMLORGESREKVKMLR 571				
Qy	642 LLSGRADYTKPAKCOBEEANSSSCTYQYQNLQQQQQPPPPPLITSPINSTLN 701				
Db	572 LLSGRADYTKPAKCOBEEANSSSCTYQYQNLQQQQQPPPPPLITSPINSTLN 604				
Qy	702 SIGTSTGGVGQGINFGPNGNNTNTGSPPPQQQQHGIITKNSPKPDRPR 761				
Db	605 -----ANLTSSSPNPKYKMQLQMN-----TQR 628				
Qy	762 ADIALYQDQHQTMTTQMLNAEKTFPSRSHSATQLOVEFNSPKHTEFSKTI 821				
Db	629 GDIALYLGQDQHQTMTTQMLNAEKTFPSRSHSATQLOVEFNSPKHTEFSKTI 688				
Qy	822 GTMLVPLNLUVLTGEGTMLGWFVGVGLPITIGSTPFAOWLKK-----	876			
Db	689 GTVLBLNLUVLTGEGTMLGWFVGVGLPITIGSTPFAOWLKK-----	747			
Qy	877 -LNNSTB-----GQNGNRPIPHSSRSRSRSLGURKRG--NRSITSPNK 920				
Db	748 TUNEAESGAKSILSSFPKRNEDKY-----SRGRSKYGGPSNRSIASLPSK 798				
Qy	921 Y 921				
Db	799 Y 799				
RESULT 5					
ID ALR1_YEAST	STANDARD	PRT;	859 AA.		

DR	Ensembl; YOL130W; Saccharomyces cerevisiae.	Db	721 YLAQLQVESFSNNKVTLEMIGKVMTLPLVNTLGFGMVKVPGF-NSSIAWFGI 779
DR	GenomeReviews; Y13140_GR; YOL130W.	DR	
DR	SGD; S00005490; ALR2.	DR	
DR	BioCyc; SCER-S28-01; SCER-S28-01-005413-MONOMER; -.	DR	
DR	LinkHub; 008269; -.	DR	
GO	GO:005886; C:plasma membrane; IMP.	GO	
GO	GO:0015082; F:di-, tri-valent inorganic cation transport; IDA.	GO	
GO	GO:0015693; F: magnesium ion transport; IDA.	GO	
InterPro	IPR012543; Mg+-transptC0RA.	InterPro	
PFAM	PF01544; CorA; 1.	PFAM	
RW	Complete proteome; Magnesium; Membrane; Transport; CHAIN 1 859	RW	
FT	Magnesium transporter ALR1.	FT	
FT	744 764 Potential.	FT	
FT	773 793 Potential.	FT	
FT	13 13 N -> Y (in Ref. 1).	FT	
SQ	SEQUENCE 859 AA; 95869 MW; 6DA4ACAT0EF2A633 CRC64;	SQ	
Query	Query Match Similarity 34.9%; Score 1680; DB 1; Length 859; Best Local Similarity 41.5%; Pred. No. 5-87; Matches 258; Indels 176; Gaps 31; Matches 405; Conservative 137; Mismatches 258; Insert 57	Query	
Db	1 MSSSEYYQNSTNPIRSDEBVLDDRRNQTNDCAISSEDBLEL--KSELESEWKS 57	Db	
Db	1 MSSSEYYQNSTNPIRSDEBVLDDRRNQTNDCAISSEDBLEL--KSELESEWKS 57	Db	
Query	58 ERQQQHQPITDSNAKPLTRKSSSIKKK---NLFQDRTINPMSLGGDDTINSQH 112	Query	
Db	50 QPPTLNKNSLAKSTPKSP1PKQSKATRYNSHDVGSPRSGRNDPDEQGQDDEVY-AHH 108	Db	
Query	113 KOR---NYNMSSL---EKFDFYLKDNTDINSNNHT---HLA1PIP1PIP1TPI 156	Query	
Db	109 QIRASAILTSARPERLAHSNPHQRYVSESNI-----HTTPKGVGTQDYTMWSSTA 161	Db	
Query	157 TNAKNSRKSQLENLPLPIKKTTGRNNNNNFENDLVSPMTKHTKTDSEDTNT-----210	Query	
Db	162 SSGNKs-KIASSASSAPTKYR-----KSLVSPVLEPHNSKSDPSKLAPEKK 210	Db	
Query	211 ---SITANMFKIGCATPLTGTCCTATA---TAAAGRGRPRSSDADSHASRSQ 264	Query	
Db	211 RTYSITSAHSSTN-PAVLTSTSQSDDADDTLERKPVMTNTRAFFEDV-SQASRSQ 268	Db	
Query	265 ETEEDVCPMPGD-HIRUNGIDPIDEFIREEEAYLOKOMIAK-NILRIDEFLNSK 322	Query	
Db	269 ETEEDVCPMPGD-HIRUNGIDPIDEFIREEEAYLOKOMIAK-NILRIDEFLNSK 322	Db	
Query	323 -----NNTTSGASRHYPYTHHSNNNNKNGDGGSSMAALKYTPKNUKLUSTRREETH 376	Query	
Db	326 DIGFTSNTSLOSS-----AALKTPR---VSQGESESTN 359	Db	
Query	377 ENSSSEREYELKIKTQOPPYKKYDOLSLTSSTSSTSGSGSSGVKFGG-ARISDGINGGSL 435	Query	
Db	360 EP-----ETHEKEDEHEK1IPS1LPG1SPKRNVEEEN1PSD1PAC1STQGTPDI 414	Db	
Query	436 PDRPLSPSESESETHAD1PS1LVS1PSQYRD1FRGEETWM1DC1TCPDSEKMLAKAF 495	Query	
Db	415 PNRSFPCSESEDETHAD1PS1LVSQCY1ELFRGEETWM1DC1SCP1DEMRC1AKAF 474	Db	
Query	496 GHPIIATAEDINQKOTREKVEYKFSKYYCPCPTEAKESDYLEKPLPNXVYVPHDQ1TF 555	Query	
Db	475 GHPIIATAEDINQKOTREKVEYKFSKYYCPCPTEAKESDYLEKPLPNXVYVPHDQ1TF 534	Db	
Query	556 HFSPISHPANTVRVRLQRD1VDSADM1CYAL1DE1TDSFAVIG1YTEDATEDAYF 615	Query	
Db	535 HGP1SHCANTRVRLQRD1VNSM1CYAL1DD1TDSFAVIQS1YTEDATEDSYF 594	Db	
Query	616 TARD1F5SMQ1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 675	Query	
Db	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	Db	
Query	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTNTG 735	Query	
Db	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	Db	
FT	TARD1F5SMQ1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 675	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT	644 -----GPAUTSQ1IANLQARDINA-----SH1KNNNSSTTPNRY 678	FT	
FT	595 MARD1DFAAM1Q1GESEKRYTMLLGLAD1KRF1ARKCDEANSSGG1QFQ1NQ 643	FT	
FT	676 QQQQQAPPPIPITSP1NSTLNLNSLG1STGCGVGGVGGINGPNTGNTNTG 735	FT	
FT			

Best Local Similarity	40.9%	Pred.	No. 3.1e-86;
Matches	401;	Conservative	126;
Indels	221;	Gaps	35;
Oy	39 DSEDBLELKSLKSELESEV----KEKQQQHQQELTSNAKPLTRKGSSIIKKSNLTDKD 93	Index	233;
Oy	9 DSSDLPLRSKSYDNTA-SMKTGKPKLENTRQ---YSDAQPTRHE---ALALKDDETKDS 62	GN	OS
Db	94 RITNPMSLGGDDTNSKKNRY---NMS-SLRDFYLKONTDNSTNNTHLAIPF 148	OC	Eukarya; Fungi; Ascomycota; Saccharomyctaceae; Saccharomyces;
Oy	63 R--HRPSSNGR----NSSEVGTVETRNNTSQRDF--EGAAEAEATYKQLASFAI- 114	NCBI_TaxID	33169;
Db	149 IPIPTITNAKSR-----RKSOLENLPIKKK-----TYGRNNNNFEND 191	RN	[1]
Oy	115 -----ISSNARPSRLAKSETHQ-QINVESTSASLPKAALERGHTPALAGTSNNRCN 168	RN	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
Db	192 LVSPTWKTIN-----PSEDINTT-----STTAHNLFGICATL 226	RN	STRAIN=ATCC 10855; DOI=10.1126/science.1095781;
Oy	169 EASSARTPSARKASLVSIAPIETSAESEGTIKQKULKRRTYSTLTHS--SYNPTTL 226	RN	RA
Db	227 GYGCTTATAATATAAG-----RPRSSSTDSEADSHASRSQETEEDVCPMTG- 276	DR	Dietrich S., Voegeli S., Brachat S., Lech A., Gates K., Steiner S., Mohr C., Poehlmann R., Luedi P., Choi S., Wing R.A., Platier A., Gaffney T.D., Philippens P.;
Oy	227 -----UTRASOKSDMNDTPRKPLMRDSEVPKSET-SQASRDQETEEDVCPMPFL 280	DR	"The Ashbya gossypii genome as a tool for mapping the ancient Saccharomyces cerevisiae genome." ; Science 304:304-307(2004).
Db	277 DILRYNGADPDEIREBREAYLQKQMLRDPKLNTRIDEFQNLNNNTGASHH---P 333	DR	Copyright by the UniProt Consortium, see http://www.uniprot.org/terms
Oy	281 LITRNGYDPELEYA-----QISNAENFLNLAMHQHQSRT 318	DR	Distributed under the Creative Commons Attribution-NoDerivs License
Db	334 YHHHSNNNNKNGDGSSSWAALKTPKNNLKKTLSRPFTTHNSSEELYLKTKQ 393	CC	CC
Oy	319 YHHDQDOPNTNSASTSOSS-AAKYTP-----ISR-----TLEKNCSTNEYAVSEANNES 368	CC	CC
Db	394 PPKYDOLISSTSSGSGQVKGARTDGNG-----SLPD 437	CC	CC
Oy	369 --VRDDDEPDL-----HEDVTPGANKTEGEKEEGNDSSYSSAYTTLQNTYQPS 415	CC	CC
Db	438 RPSLPHSESETHAHPDPSLVSQVRDPLRNGBETWMDOCTCPSEMMOMLAKAFG 497	CC	CC
Oy	416 RSPFRSEDEVRAASDPSLSSGQFPTFLFKGQDFTWDLQCDPDEBMCIAKNTG 475	CC	CC
Db	498 RPLTAEDTRMQETRKEVPLKPSYFCPTFEAKDSESDELYEPINIVYTFPHDGLTYPHP 557	CC	CC
Oy	476 HPLTDEDRMQETRKEVPLKPSYFCPTFHENDKESENYLFINIVYTFPSGULTHF 535	CC	CC
Db	558 SPISHPANVRVRVQLRQDYYDVSADMWCAYALIDETDGRAPHVHGKEYEADIEAVFTA 617	CC	CC
Oy	536 DPISPCANVRVRVQLRDYVSNSDWLYALIDDTSPAPVQISYEADSIDSUVF 595	CC	CC
Db	618 RUDPDSHSLQIGRSRKVTLKLGRADVKMFKAQCEBANSCTYORQ---TNYL 674	CC	CC
Oy	596 RUMDFAAMLQIGRSRKTMKLKLGRADVKMFKAQCDETNGIGEVLKQSONMMNL 655	CC	CC
Db	675 QQQQQAPPNPWITSPINSTNLSQTGGVGVGINFGENPGCQNTNTNTT 734	CC	CC
Oy	656 QNEQP-----NVQ-----NNSN-----668	CC	CC
Db	735 GSPSPPOOOQHGNTKNSPKLIPDARPAFDIALYGDQDHTIMTFONLLAYEKIFSRSHS 794	CC	CC
Oy	669 -----NQISLNSN-SYNOTSNSPKTMLKLGRADVKMFKAQCDETNGIGEVLKQSONMMNL 718	CC	CC
Db	795 NYLAQLOQESNSPKNTENFSKTLGTLVNLULVLFQANTVPEOEGTNLWTGF 854	CC	CC
Oy	719 NYLAQLOQESNSPKNTENFSKTLGTLVNLULVLFQANTVPEOEGTNLWTGF 777	CC	CC
Db	855 IVGVLIPTIGSPFAOMWLKLNST-----ECONGNRP-----PNHSSRSI 900	CC	CC
Oy	778 ILGULLLLAVSWPLASWIKKDPPTALNEAAGSACKSVISSFLPKRDKRPNDSKGN 837	CC	CC
Db	901 RSLGLKHKGGKNSLISPPN 921	CC	CC
Oy	838 ARVGYRR----KSTVSLPSRY 854	CC	CC
Db	RESULT 7	CC	CC
OYAU1_ASHGO	PRELIMINARY;	PRY;	799 AA.
ID Q7SAU1_ASHGO			
AC Q7SAU1;			
DT 05-JUL-2004,	integrated into UniProtKB/TrEMBL.		
DT 05-JUL-2004,	sequence version 1.		
DT 07-FEB-2006,	entry version 16.		

DE ADL171Wp.	GN Order=LocustNames=ADL171W;	OS Ashbya gossypii (Yeast) (Bremothecium gossypii).
OY	9 DSSDLPLRSKSYDNTA-SMKTGKPKLENTRQ---YSDAQPTRHE---ALALKDDETKDS 62	OC Saccharomyctaceae; Saccharomyces;
DB	94 RITNPMSLGGDDTNSKKNRY---NMS-SLRDFYLKONTDNSTNNTHLAIPF 148	OX Eremothecium.
OY	63 R--HRPSSNGR----NSSEVGTVETRNNTSQRDF--EGAAEAEATYKQLASFAI- 114	NCBI_TaxID
DB	149 IPIPTITNAKSR-----RKSOLENLPIKKK-----TYGRNNNNFEND 191	RN [1]
OY	115 -----ISSNARPSRLAKSETHQ-QINVESTSASLPKAALERGHTPALAGTSNNRCN 168	RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
DB	192 LVSPTWKTIN-----PSEDINTT-----STTAHNLFGICATL 226	RN STRAIN=ATCC 10855; DOI=10.1126/science.1095781;
OY	169 EASSARTPSARKASLVSIAPIETSAESEGTIKQKULKRRTYSTLTHS--SYNPTTL 226	RN DR
DB	227 GYGCTTATAATATAAG-----RPRSSSTDSEADSHASRSQETEEDVCPMTG- 276	DR DR
OY	227 -----UTRASOKSDMNDTPRKPLMRDSEVPKSET-SQASRDQETEEDVCPMPFL 280	DR DR
DB	277 DILRYNGADPDEIREBREAYLQKQMLRDPKLNTRIDEFQNLNNNTGASHH---P 333	DR DR
OY	281 LITRNGYDPELEYA-----QISNAENFLNLAMHQHQSRT 318	DR DR
DB	334 YHHHSNNNNKNGDGSSSWAALKTPKNNLKKTLSRPFTTHNSSEELYLKTKQ 393	DR DR
OY	319 YHHDQDOPNTNSASTSOSS-AAKYTP-----ISR-----TLEKNCSTNEYAVSEANNES 368	DR DR
DB	394 PPKYDOLISSTSSGSGQVKGARTDGNG-----SLPD 437	DR DR
OY	369 --VRDDDEPDL-----HEDVTPGANKTEGEKEEGNDSSYSSAYTTLQNTYQPS 415	DR DR
DB	438 RPSLPHSESETHAHPDPSLVSQVRDPLRNGBETWMDOCTCPSEMMOMLAKAFG 497	DR DR
OY	416 RSPFRSEDEVRAASDPSLSSGQFPTFLFKGQDFTWDLQCDPDEBMCIAKNTG 475	DR DR
DB	498 RPLTAEDTRMQETRKEVPLKPSYFCPTFEAKDSESDELYEPINIVYTFPHDGLTYPHP 557	DR DR
OY	476 HPLTDEDRMQETRKEVPLKPSYFCPTFHENDKESENYLFINIVYTFPSGULTHF 535	DR DR
DB	558 SPISHPANVRVRVQLRQDYYDVSADMWCAYALIDETDGRAPHVHGKEYEADIEAVFTA 617	DR DR
OY	536 DPISPCANVRVRVQLRDYVSNSDWLYALIDDTSPAPVQISYEADSIDSUVF 595	DR DR
DB	618 RUDPDSHSLQIGRSRKVTLKLGRADVKMFKAQCEBANSCTYORQ---TNYL 674	DR DR
OY	596 RUMDFAAMLQIGRSRKTMKLKLGRADVKMFKAQCDETNGIGEVLKQSONMMNL 655	DR DR
DB	675 QQQQQAPPNPWITSPINSTNLSQTGGVGVGINFGENPGCQNTNTNTT 734	DR DR
OY	656 QNEQP-----NVQ-----NNSN-----668	DR DR
DB	735 GSPSPPOOOQHGNTKNSPKLIPDARPAFDIALYGDQDHTIMTFONLLAYEKIFSRSHS 794	DR DR
OY	669 -----NQISLNSN-SYNOTSNSPKTMLKLGRADVKMFKAQCDETNGIGEVLKQSONMMNL 718	DR DR
DB	795 NYLAQLOQESNSPKNTENFSKTLGTLVNLULVLFQANTVPEOEGTNLWTGF 854	DR DR
OY	719 NYLAQLOQESNSPKNTENFSKTLGTLVNLULVLFQANTVPEOEGTNLWTGF 777	DR DR
DB	855 IVGVLIPTIGSPFAOMWLKLNST-----ECONGNRP-----PNHSSRSI 900	DR DR
OY	778 ILGULLLLAVSWPLASWIKKDPPTALNEAAGSACKSVISSFLPKRDKRPNDSKGN 837	DR DR
DB	901 RSLGLKHKGGKNSLISPPN 921	DR DR
OY	838 ARVGYRR----KSTVSLPSRY 854	DR DR
DB	RESULT 7	DR DR
OYAU1_ASHGO	PRELIMINARY;	PRY;
ID Q7SAU1_ASHGO		
AC Q7SAU1;		
DT 05-JUL-2004,	integrated into UniProtKB/TrEMBL.	
DT 05-JUL-2004,	sequence version 1.	
DT 07-FEB-2006,	entry version 16.	
DE ADL171W.	GN Order=LocustNames=ADL171W;	OS Ashbya gossypii (Yeast) (Bremothecium gossypii).
OY	9 DSSDLPLRSKSYDNTA-SMKTGKPKLENTRQ---YSDAQPTRHE---ALALKDDETKDS 62	OC Saccharomyctaceae; Saccharomyces;
DB	94 RITNPMSLGGDDTNSKKNRY---NMS-SLRDFYLKONTDNSTNNTHLAIPF 148	OX Eremothecium.
OY	63 R--HRPSSNGR----NSSEVGTVETRNNTSQRDF--EGAAEAEATYKQLASFAI- 114	NCBI_TaxID
DB	149 IPIPTITNAKSR-----RKSOLENLPIKKK-----TYGRNNNNFEND 191	RN [1]
OY	115 -----ISSNARPSRLAKSETHQ-QINVESTSASLPKAALERGHTPALAGTSNNRCN 168	RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
DB	192 LVSPTWKTIN-----PSEDINTT-----STTAHNLFGICATL 226	RN STRAIN=ATCC 10855; DOI=10.1126/science.1095781;
OY	169 EASSARTPSARKASLVSIAPIETSAESEGTIKQKULKRRTYSTLTHS--SYNPTTL 226	RN DR
DB	227 GYGCTTATAATATAAG-----RPRSSSTDSEADSHASRSQETEEDVCPMTG- 276	DR DR
OY	227 -----UTRASOKSDMNDTPRKPLMRDSEVPKSET-SQASRDQETEEDVCPMPFL 280	DR DR
DB	277 DILRYNGADPDEIREBREAYLQKQMLRDPKLNTRIDEFQNLNNNTGASHH---P 333	DR DR
OY	281 LITRNGYDPELEYA-----QISNAENFLNLAMHQHQSRT 318	DR DR
DB	334 YHHHSNNNNKNGDGSSSWAALKTPKNNLKKTLSRPFTTHNSSEELYLKTKQ 393	DR DR
OY	319 YHHDQDOPNTNSASTSOSS-AAKYTP-----ISR-----TLEKNCSTNEYAVSEANNES 368	DR DR
DB	394 PPKYDOLISSTSSGSGQVKGARTDGNG-----SLPD 437	DR DR
OY	369 --VRDDDEPDL-----HEDVTPGANKTEGEKEEGNDSSYSSAYTTLQNTYQPS 415	DR DR
DB	438 RPSLPHSESETHAHPDPSLVSQVRDPLRNGBETWMDOCTCPSEMMOMLAKAFG 497	DR DR
OY	416 RSPFRSEDEVRAASDPSLSSGQFPTFLFKGQDFTWDLQCDPDEBMCIAKNTG 475	DR DR
DB	498 RPLTAEDTRMQETRKEVPLKPSYFCPTFEAKDSESDELYEPINIVYTFPHDGLTYPHP 557	DR DR
OY	476 HPLTDEDRMQETRKEVPLKPSYFCPTFHENDKESENYLFINIVYTFPSGULTHF 535	DR DR
DB	558 SPISHPANVRVRVQLRQDYYDVSADMWCAYALIDETDGRAPHVHGKEYEADIEAVFTA 617	DR DR
OY	536 DPISPCANVRVRVQLRDYVSNSDWLYALIDDTSPAPVQISYEADSIDSUVF 595	DR DR
DB	618 RUDPDSHSLQIGRSRKVTLKLGRADVKMFKAQCEBANSCTYORQ---TNYL 674	DR DR
OY	596 RUMDFAAMLQIGRSRKTMKLKLGRADVKMFKAQCDETNGIGEVLKQSONMMNL 655	DR DR
DB	675 QQQQQAPPNPWITSPINSTNLSQTGGVGVGINFGENPGCQNTNTNTT 734	DR DR
OY	656 QNEQP-----NVQ-----NNSN-----668	DR DR
DB	735 GSPSPPOOOQHGNTKNSPKLIPDARPAFDIALYGDQDHTIMTFONLLAYEKIFSRSHS 794	DR DR
OY	669 -----NQISLNSN-SYNOTSNSPKTMLKLGRADVKMFKAQCDETNGIGEVLKQSONMMNL 718	DR DR
DB	795 NYLAQLOQESNSPKNTENFSKTLGTLVNLULVLFQANTVPEOEGTNLWTGF 854	DR DR
OY	719 NYLAQLOQESNSPKNTENFSKTLGTLVNLULVLFQANTVPEOEGTNLWTGF 777	DR DR
DB	855 IVGVLIPTIGSPFAOMWLKLNST-----ECONGNRP-----PNHSSRSI 900	DR DR
OY	778 ILGULLLLAVSWPLASWIKKDPPTALNEAAGSACKSVISSFLPKRDKRPNDSKGN 837	DR DR
DB	901 RSLGLKHKGGKNSLISPPN 921	DR DR
OY	838 ARVGYRR----KSTVSLPSRY 854	DR DR
DB	RESULT 7	DR DR
OYAU1_ASHGO	PRELIMINARY;	PRY;
ID Q7SAU1_ASHGO		
AC Q7SAU1;		
DT 05-JUL-2004,	integrated into UniProtKB/TrEMBL.	
DT 05-JUL-2004,	sequence version 1.	
DT 07-FEB-2006,	entry version 16.	
DE ADL171W.	GN Order=LocustNames=ADL171W;	OS Ashbya gossypii (Yeast) (Bremothecium gossypii).
OY	9 DSSDLPLRSKSYDNTA-SMKTGKPKLENTRQ---YSDAQPTRHE---ALALKDDETKDS 62	OC Saccharomyctaceae; Saccharomyces;
DB	94 RITNPMSLGGDDTNSKKNRY---NMS-SLRDFYLKONTDNSTNNTHLAIPF 148	OX Eremothecium.
OY	63 R--HRPSSNGR----NSSEVGTVETRNNTSQRDF--EGAAEAEATYKQLASFAI- 114	NCBI_TaxID
DB	149 IPIPTITNAKSR-----RKSOLENLPIKKK-----TYGRNNNNFEND 191	RN [1]
OY	115 -----ISSNARPSRLAKSETHQ-QINVESTSASLPKAALERGHTPALAGTSNNRCN 168	RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
DB	192 LVSPTWKTIN-----PSEDINTT-----STTAHNLFGICATL 226	RN STRAIN=ATCC 10855; DOI=10.1126/science.1095781;
OY	169 EASSARTPSARKASLVSIAPIETSAESEGTIKQKULKRRTYSTLTHS--SYNPTTL 226	RN DR
DB	227 GYGCTTATAATATAAG-----RPRSSSTDSEADSHASRSQETEEDVCPMTG- 276	DR DR
OY	227 -----UTRASOKSDMNDTPRKPLMRDSEVPKSET-SQASRDQETEEDVCPMPFL 280	DR DR
DB	277 DILRYNGADPDEIREBREAYLQKQMLRDPKLNTRIDEFQNLNNNTGASHH---P 333	DR DR
OY	281 LITRNGYDPELEYA-----QISNAENFLNLAMHQHQSRT 318	DR DR
DB	334 YHHHSNNNNKNGDGSSSWAALKTPKNNLKKTLSRPFTTHNSSEELYLKTKQ 393	DR DR
OY	319 YHHDQDOPNTNSASTSOSS-AAKYTP-----ISR-----TLEKNCSTNEYAVSEANNES 368	DR DR
DB	394 PPKYDOLISSTSSGSGQVKGARTDGNG-----SLPD 437	DR DR
OY	369 --VRDDDEPDL-----HEDVTPGANKTEGEKEEGNDSSYSSAYTTLQNTYQPS 415	DR DR
DB	438 RPSLPHSESETHAHPDPSLVSQVRDPLRNGBETWMDOCTCPSEMMOMLAKAFG 497	DR DR
OY	416 RSPFRSEDEVRAASDPSLSSGQFPTFLFKGQDFTWDLQCDPDEBMCIAKNTG 475	DR DR
DB	498 RPLTAEDTRMQETRKEVPLKPSYFCPTFEAKDSESDELYEPINIVYTFPHDGLTYPHP 557	DR DR
OY	476 HPLTDEDRMQETRKEVPLKPSYFCPTFHENDKESENYLFINIVYTFPSGULTHF 535	DR DR
DB	558 SPISHPANVRVRVQLRQDYYDVSADMWCAYALIDETDGRAPHVHGKEYEADIEAVFTA 617	DR DR
OY	536 DPISPCANVRVRVQLRDYVSNSDWLYALIDDTSPAPVQISYEADSIDSUVF 595	DR DR
DB	618 RUDPDSHSLQIGRSRKVTLKLGRADVKMFKAQCEBANSCTYORQ---TNYL 674	DR DR
OY	596 RUMDFAAMLQIGRSRKTMKLKLGRADVKMFKAQCDETNGIGEVLKQSONMMNL 655	DR DR
DB	675 QQQQQAPPNPWITSPINSTNLSQTGGVGVGINFGENPGCQNTNTNTT 734	DR DR
OY	656 QNEQP-----NVQ-----NNSN-----668	DR DR
DB	735 GSPSPPOOOQHGNTKNSPKLIPDARPAFDIALYGDQDHTIMTFONLLAYEKIFSRSHS 794	DR DR
OY	669 -----NQISLNSN-SYNOTSNSPKTMLKLGRADVKMFKAQCDETNGIGEVLKQSONMMNL 718	DR DR
DB	795 NYLAQLOQESNSPKNTENFSKTLGTLVNLULVLFQANTVPEOEGTNLWTGF 854	DR DR
OY	719 NYLAQLOQESNSPKNTENFSKTLGTLVNLULVLFQANTVPEOEGTNLWTGF 777	DR DR
DB	855 IVGVLIPTIGSPFAOMWLKLNST-----ECONGNRP-----PNHSSRSI 900	DR DR
OY	778 ILGULLLLAVSWPLASWIKKDPPTALNEAAGSACKSVISSFLPKRDKRPNDSKGN 837	DR DR
DB	901 RSLGLKHKGGKNSLISPPN 921	DR DR
OY	838 ARVGYRR----KSTVSLPSRY 854	DR DR
DB	RESULT 7	DR DR
OYAU1_ASHGO	PRELIMINARY;	PRY;
ID Q7SAU1_ASHGO		
AC Q7SAU1;		
DT 05-JUL-2004,	integrated into UniProtKB/TrEMBL.	
DT 05-JUL-2004,	sequence version 1.	
DT 07-FEB-2006,	entry version 16.	
DE ADL171W.	GN Order=LocustNames=ADL171W;	OS Ashbya gossypii (Yeast) (Bremothecium gossypii).
OY	9 DSSDLPLRSKSYDNTA-SMKTGKPKLENTRQ---YSDAQPTRHE---ALALKDDETKDS 62	OC Saccharomyctaceae; Saccharomyces;
DB	94 RITNPMSLGGDDTNSKKNRY---NMS-SLRDFYLKONTDNSTNNTHLAIPF 148	OX Eremothecium.
OY	63 R--HRPSSNGR----NSSEVGTVETRNNTSQRDF--EGAAEAEATYKQLASFAI- 114	NCBI_TaxID
DB	149 IPIPTITNAKSR-----RKSOLENLPIKKK-----TYGRNNNNFEND 191	RN [1]
OY	115 -----ISSNARPSRLAKSETHQ-QINVESTSASLPKAALERGHTPALAGTSNNRCN 168	RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
DB	192 LVSPTWKTIN-----PSEDINTT-----STTAHNLFGICATL 226	RN STRAIN=ATCC 10855; DOI=10.1126/science.1095781;
OY	169 EASSARTPSARKASLVSIAPIETSAESEGTIKQKULKRRTYSTLTHS--SYNPTTL 226	RN DR
DB	227 GYGCTTATAATATAAG-----RPRSSSTDSEADSHASRSQETEEDVCPMTG- 276	DR DR
OY	227 -----UTRASOKSDMNDTPRKPLMRDSEVPKSET-SQASRDQETEEDVCPMPFL 280	DR DR
DB	277 DILRYNGADPDEIREBREAYLQKQMLRDPKLNTRIDEFQNLNNNTGASHH---P 333	DR DR
OY	281 LITRNGYDPELEYA-----QISNAENFLNLAMHQHQSRT 318	DR DR
DB	334 YHHHSNNNNKNGDGSSSWAALKTPKNNLKKTLSRPFTTHNSSEELYLKTKQ 393	DR DR
OY	319 YHHDQDOPNTNSASTSOSS-AAKYTP-----ISR-----TLEKNCSTNEYAVSEANNES 368	DR DR
DB	394 PPKYDOLISSTSSGSGQVKGARTDGNG-----SLPD 437	DR DR
OY	369 --VRDDDEPDL-----HEDVTPGANKTEGEKEEGNDSSYSSAYTTLQNTYQPS 415	DR DR
DB	438 RPSLPHSESETHAHPDPSLVSQVRDPLRNGBETWMDOCTCPSEMMOMLAKAFG 497	DR DR
OY	416 RSPFRSEDEVRAASDPSLSSGQFPTFLFKGQDFTWDLQCDPDEBMCIAKNTG 475	DR DR
DB	498 RPLTAEDTRMQETRKEVPLKPSYFCPTFEAKDSESDELYEPINIVYTFPHDGLTYPHP 557	DR DR
OY	476 HPLTDEDRMQETRKEVPLKPSYFCPTFHENDKESENYLFINIVYTFPSGULTHF 535	DR DR
DB	558 SPISHPANVRVRVQLRQDYYDVSADMWCAYALIDETDGRAPHVHGKEYEADIEAVFTA 617	DR DR
OY	536 DPISPCANVRVRVQLRDYVSNSDWLYALIDDTSPAPVQISYEADSIDSUVF 595	DR DR
DB	618 RUDPDSHSLQIGRSRKVTLKLGRADVKMFKAQCEBANSCTYORQ---TNYL 674	DR DR
OY	596 RUMDFAAMLQIGRSRKTMKLKLGRADVKMFKAQCDETNGIGEVLKQSONMMNL 655	DR DR
DB	675 QQQQQAPPNPWITSPINSTNLSQTGGVGVGINFGENPGCQNTNTNTT 734	DR DR
OY	656 QNEQP-----NVQ-----NNSN-----668	DR DR
DB	735 GSPSPPOOOQHGNTKNSPKLIPDARPAFDIALYGDQDHTIMTFONLLAYEKIFSRSHS 794	DR DR
OY	669 -----NQISLNSN-SYNOTSNSPKTMLKLGRADVKMFKAQCDETNGIGEVLKQSONMMNL 718	DR DR
DB	795 NYLAQLOQESNSPKNTENFSKTLGTLVNLULVLFQANTVPEOEGTNLWTGF	

Matches	310:	Conservative 100:	Mismatches 247:	Indels 220:	Gaps 22:
RC	STRAIN=972 h-;				
RC	MEDLINE=200890407; Published=10620777;				
RX	DOI=10.1002/(SICI)1097-0061(20000115)16:1<71::AID-YEAS05>3.0.CO;2-5;				
RX	Machida M., Yamazaki S., Kunihiro S., Tanaka T., Kusuda N., Jinno K.,				
RA	Hakawa Y., Yamamoto J., Yamamoto K., Ogura K., Kudo H., Kikuchi A., Nagai Y.,				
RA	Yanagida M., Aoki K., Ogura K., Kudo H., Kikuchi H., Zhang M.Q.,				
RA	A 18 kb segment containing the cdc2 gene from the left arm of fission yeast chromosome II: sequence analysis and characterization of the genomic DNA and cDNAs encoded on the segment.;				
RT	Yeast 16:71-80 (2000).				
RN	[2]				
RP	NUCLEOTIDE SEQUENCE.				
RC	STRAIN=972 h-;				
RC	Director-General of Biotechnology Center, Kusuda N., Machida M.;				
RA	Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.				
RL	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms				
CC	Distributed under the Creative Commons Attribution-NoDerivs License				
CC	EMBL: AB004539; BAA21448_1; -: Genomic DNA.				
DR	GO:0016020; C:membrane; IEA.				
DR	GO:0046873; P:metal ion transport; IEA.				
DR	GO:0030001; P:metal ion transport; IEA.				
DR	InterPro: IPR025253; Mg+_transptCora.				
DR	Pfam: PF0144; Cora: 1.				
FT	NON_TER 1 1				
SQ	SEQUENCE 499 AA: BBC3715E9B6D2C4C CRC64:				
Query	7 TESTTQPSDETFVTPRKLPSSTSSADENNEVDQVSEDTNTSTANTAKPLGIGATTLGV-GTCATTATAATAA 242	23.7%	Score 1141.5;	DB 2;	Length 499;
Query	184 NSNFENDDLVSPMTKMTNSDIDTNTSTANTAKPLGIGATTLGV-GTCATTATAATAA 242	20.9%	Best Local Similarity	Pred. No. 1.2e-56;	
Query	299 EAYLKQMIANILRIDEQNLKSANTTSGASH-----PYHHSNNNNKANGGGGSS 353	20.9%	Matches 267;	Conservative 60;	Gaps 13
Query	267 EDSPTMVPRPSPATKSACAAVDSPHTTIPPPYQSDTDLTEUPTKSTAQSIEFTQP 326	20.9%	Matches 267;	Mismatches 60;	
Query	385 ---LOKQ-----RQNDTHFRPSQYSTCKPEFHAWDLSPHDND-----420	20.9%			
Query	243 GRRPSS1----DSEADSHASRSQTEEDVCPMVGDHIRNGIDFDEDFREERE 298	20.9%			
Query	327 RUKPSHIANNEDEVDQVSEDTNTSTANTAKPLGIGATTLGV-GTCATTATAATAA 384	20.9%			
Query	421 -----PSSL-----HSNNAAKAEPVPRSS--YSGSGRSVTSMSDS--- 456	20.9%			
Query	414 SGSGVVKPGARISDGINGSLPDRFLSFISSEBETHADIPSIVSPG-QSVRLPFRNG 472	20.9%			
Query	457 -----PTRFPSNSQNETHAVSLDLDGATSFRSIFCFPCE 494	20.9%			
Query	473 RETWNLDCTCPDSSEMOKMLAKAFGHLUTLAEDIRMOETREKVELPKSKYXPFHTEARK 532	20.9%			
Query	495 KGWWLDCLDPTDIEVRLSKAFIHLTTEDIVQEARKEVLFSSTYFCRSEQDP 554	20.9%			
Query	533 ESDPBDPAINVIVPHDG ITTFAPSPISPLSNSVAVRVRVRLDQYDVSADWLYCALIDII 592	20.9%			
Query	555 ELANTYLPENATIVVPRGCLTFIFPSATHPSAVERAROLRQYRVSSDWLYCALIDII 614	20.9%			
Query	593 TDGFAPVPHGELEYADATDVAETARDPSMQLQIGEERRVVMTLMLLSKEDADITION 652	20.9%			
Query	615 TDAFPFLRGTEETEAIECSVLYGRSSEDDSMILRRGECRKTAGMPRLTKADWIKM 674	20.9%			
Query	653 FAKRCQEAEASSGGTYQORXNLQOCQQOAPPNNPNTSPINSTLNLSGTSTGGVG 712	20.9%			
Query	675 LAKRNKRWTIA-----686	20.9%			
Query	713 VGGINFGPNPQNTNTNTGSPSPQQQHQGITNKSFPPDARPDATLYLGDIQ 772	20.9%			
Query	687 -----PTG-----EGLTGLDQ 699	20.9%			
Query	773 DHITMFONLAYERKIFSRSHSNYLAQVBSFNINNKTEPEFSKITLEGPMLVPAULVT 832	20.9%			
Query	700 DHLVMTSNLSQENILSRSHSNYLAQVTSNEENSRTINGAQLGKLTGTLVPMNLVT 759	20.9%			
Query	833 GLPGRNVPVUGPSEGTTGCMWPGTIVGULI-PIIGSFI 868	20.9%			
Query	760 GLPGRNVPVUGDNTLNLAFFGILVULGSIAIGWII 796	20.9%			
Db	RESULTS 11				
Db	71LWDB SCHPO Q7LWDB_SCHPO PRELIMINARY; PRTE: 499 AA.				
Db	07-JUL-2004, integrated into UniProtKB/TREMBL.				
Db	05-JUL-2004, sequence version 1.				
Db	07-FEB-2006, entry version 8.				
Db	Probable membrane protein YOL130w (Fragment).				
Db	Name: YOL130w; Name2: YOL130w; Name3: YOL130w; Name4: YOL130w.				
Db	Schizosaccharomyces pombe (Fission yeast).				
Db	Schizosaccharomyces: Schizosaccharomyces; Schizosaccharomyces: Schizosaccharomyces; Schizosaccharomyces: Schizosaccharomyces.				
Db	NCBI TaxID:48936;				
X	NCBI LOCATOR: Sequence.				

QWTR5_ASPFU	PRELIMINARY;	PRT;	663 AA.		345	NGGCGGSSAAALKYTPKNIKTKLRSRFPETHNSSSSEIYLKTKQPPKYDQLSL 404	Qy
AC				:			Db
DD	03-JUL-2005,	Integrated into UniProtKB/TREMBL.				--DRSSIDLMDADDLTAKTADKVANVVERD-	304
DT	05-JUL-2005,	sequence version 1.					
DT	07-MAR-2006,	entry version 6.					
DB		CorA family metal ion transporter, putative.					
GS	ORFName=Af5g05830;						
OC	Aspergillus fumigatus (Sarcoryza fumigata)	Fungi: Ascomycota: Pezizomycotina: Eurotiomycetes;					
OC	Eukaryota; Trichomycetidae; mitosporic Trichocomaceae; Aspergillales; Aspergillus						
OX	NCBI_TaxID=083;						
RN	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].						
FC	STRAIN=f6291 / CBS 10135 / FSC A1100.						
FX	PubMed=16372009; DOI=10.1038/nature04332;						
RA	Nierman W.C., Pain A., Anderson N.J., Worman J.R., Kim H.S., Arroyo J., Berriman M., Abe K., Archer D.B., Bennett J.W., Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S., Farman M., Fedorova N., Fedorova N.I., Feidibiyum T.V., Fischer R., Fosker N., Fraser A., Garcia J.L., Garcia M.J., Goble A., Goldman G.H., Gonii K., Griffiths Jones S., Gwilliam R., Haas B.J., Haas H., Harris D.E., Horisuchi H., Huang J., Humphrey S., Jimenez J., Keller N., Khouri H., Kitamoto K., Kobayashi T., Konzack S., Kulkarni R., Majors W.H., May G.S., Miller B.L., Mohamoud A., Molina M., Monod M., Mouyna I., Mulligan S., Murphy L.D., O'Neill S., Paulsen I., Penalya M.A., Pertea M., Price C., Pritchard B.L., Quail M.A., Rabbinowitch E., Rawlins N., Rajandream M.A., Reichard U., Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M., Ronning C.M., Rutter S., Selzberg S.L., Sanchez M., Takeuchi M., Tekala P., Saunders D., Seeger K., Squares S., White O., Woodward J.R., Yu J.-H., Fraser C.M., Galagan J.E., Machida M., Hall N., Barrell B.G., Denning D.W., Rattray M., Asai K., Aspergillus fumigatus.*						
RL	Nature (38:1151-1156/2005).						
CC	-:- GROWTH: The sequence shown here is derived from an EMBL/Genbank/DBJ/Whole genome shotgun (WGS) entry which is preliminary data.						
CC	6 SEYYQNNTTQP1PRTSPDEVLDIIRNQITNDCAISDSEBLKSELESEVVKSKPQQH 63						
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms						
CC	Distributed under the Creative Commons Attribution-NoDerivs License						
DR	EMBL:AMPH1000003; ENR92011.; -; Genomic_DNA.						
DR	GO:0016820; C:membrane; IEA.						
DR	GO:00046873; F:metal ion transport; IEA.						
DR	GO:0030001; P:metal ion transport; IEA.						
KW	Complete proteome.						
SQ	SEQUENCE 663 AA; 74724 MW; 10CFCE563D543C CRC64;						
Query Match	22.8%; Score 1099; DB 2; Length 663;						
Best Local Similarity	34.6%; Fred. No. 4.5e-54;						
Matches	291; Conservative 96; Mismatches 245; Indels 260; Gaps 21;						
CC	4 SEYYQNNTTQP1PRTSPDEVLDIIRNQITNDCAISDSEBLKSELESEVVKSKPQQH 63						
CC	7 SEKXMC1LTFRTSPVPE---LDDHRFQL---DSPPRATATAI-SLSQNTAQAY 54						
CC	64 ROBTSMDNAKPLTRKSGSSIIKKKSNLTDKRITNPMSLSSGGDTTNSGHKNYNNMSLR 123						
CC	55 HOETFO-----RDPLS1QDALRAGLSLSDFQATID 87						
CC	124 KOFYTKD-----NTDDNSTNNH----TMALAP1P1PTP1PTITANKSREK--SQ 167						
CC	88 DBRSCKDINTLGRRTSPDNQNPYHPTWSTQEL-----ANMRSSESSA 134						
CC	168 LENLPFLIKKTTGRNNSNNFENDLYSPMTKRMNTDITNTSTANFKLUGATLG 227						
CC	135 RSSSP-----NSVAFADPRERFANTEHATEDLIA-----ILQRVTS 175						
CC	228 VGTGTTATAAAAGRPSRSRSDSEADSHASRSQSSEEDVCPMWGDHTRVNGIDF 287						
CC	176 GTTHRRPFPSMHA-TRPQGDIQEPN-----DESCVTEQPRIVIDYE 223						
CC	286 EDEPTEPEREAYLQMQMLAKNLRIDEQUNLSKUNNTGASRHP---YHHHSNNNNCRN 344						
CC	224 ELEEFVA-----LSRQMKPSTSR-----KQSLSQSGRGRVYDPLRSR 267						
DR	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms						
DR	Distributed under the Creative Commons Attribution-NoDerivs License						
DR	EMBL: CR32121; CAS78951.; -; Genomic_DNA.						
DR	GO: GO:0016030; C:membrane; IEA.						
DR	GO: GO:0046873; F:metal ion transporter activity; IEA.						

DR	GO:0030001; Pimelat ion transport; IEA.	RA	Lui A., Ma L.-J., Mabbitt R., MacDonald J., MacLean C., Major J.,
DR	InterPro; IPR01356; Homeobox.	RA	Manning J., Maru K., Matthee T., Mauceli E.,
DR	InterPro; IPR02543; Mg++_transportCora.	RA	McCarthy M., McDonough S., McHugh T., Meldrum J., Menetsky J.,
DR	Pf01544; CorA; 1.	RA	Mikhael A., Mihowa T., Mlenga V., Moru K.,
DR	PROSITE; PS00027; HOMEOROX_1; UNKNOWN_1.	RA	Nesirov J., Nihalayev A., Nguyen C., Nguyen C.,
KW	Complete proteome.	RA	Newes C., Munson G., Nguyen C.,
SEQUENCE	52281 MW; 50404800DPD76400 CRC64;	RA	Nguyen N., Nguyen T., Nicoll R., Nielsen C., Norbu C.,
SEQUENCE	460 AA;	RA	Norbu N., O'Donnell P., Okamoto O., O'Leary S., Omotoshi B.,
Qy	Best Local Similarity 21.7%; Score 1044; DB 2; Length 460;	RA	O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Picqani B.,
Matches 202; Conservative 76; Mismatches 93; Index 104; Gaps 3;	RA	Purcell S., Rachupka T., Ramamurthy U., Rameau R., Ray V., Raymond C.,	
Qy	437 DREFLPHSESEETHTADPISVSEGSODRFLNGEETWLDCTCPTDSEKMKLLAKAFG 496	RA	Retta R., Richardson S., Risse C., Rodriguez J., Rogers J., Rogov P.,
Db	77 DHPAFSTCECTTIANEPSETEDGCKAKELPAFKGTWMLCDPTDAENKLAKAFG 136	RA	Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe J., Sounez C.,
Qy	497 IHEPLAEDTRMOTERKEVLFSTYFQPHFEADKSESEDYLPINIVYLVFDGLTH 556	RA	Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sounez C.,
Db	137 IHPFLTYDQHREAKEYEMPRNTYPAFHAYEDVDESDFMFAVPPYLCPVEGVYSPH 196	RA	Spencer B., Staller J., Strange-Thomann M., Stavropoulos S.,
Qy	557 FSPISHPANVERRVRQLRQLDYDVSADMICVLTALIDTEITDGAPVHGLEYADAIEDAFT 616	RA	Stetson K., Stone C., Stone S., Stubbs M., Talavera J., Tchuinga P.,
Db	197 YSFVPICATVRRKIQIQLDHTVTPDWIHAVIDDSAPVTRIEQDGEAEVQTFD 256	RA	Tenzing P., Tesfaye S., Theodore J., Thounthasang Y., Topham K.,
Qy	617 ARDDESSLQLQIGTSRUKVMTARLISGRADYKTFPAKRCQBEANSSGGYQRQNLQO 676	RA	Tewey S., Tsamli T.S., Tsamli N., Vallee D., Vassiliev H., Venkataraman J.S., Vinson J.J., Vo A., Wade C., Wang S., Wangchuk T.,
Db	257 ARDGFNKLRLIGKARAKTUSMRHLISGRADYVGMETKVSF----- 299	RA	Wangdi T., Whittaker C., Williamson J., Wu A., Yadav S., Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zenbeck L., Zimmer A., Zody M., Lander E.S.;
Qy	677 QQQQA PPPPPPBLITSPINSTNLNLNLGTSGCCVGYGGINFGPNPFGANTNTNTCS 736	RA	*!- CAUT: The genome sequence of <i>Ustilago maydis</i> . *
Db	300 ----- 299	RA	-!- Submitted (PER-204) to the EMBL/GenBank/DBJ databases.
Qy	737 PSPRQOOQHGTWSKSPPIPDARPADIALYLGDIQHITMFQNLLAYEKIFSRSHNY 796	RA	-!- The genome sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.
Db	300 ----- GLSDRHKGHTSLYLGDIQHITWYTOSIJAHKULSLSHNY 342	CC	CC
Qy	797 LAQLOQVEPSNNKITEFSKTILGTMPLVNLTCPLGEMATVPGEGST---NLGMPF 852	CC	Copyright by the UniProt Consortium, see http://www.uniprot.org/terms
Db	343 LSOLQVOSDIAHTHRYTDTGKITYTGTTGILIPMNVTGLAFAMNRIGESTDENPGNFMF 402	CC	Distributed under the Creative Commons Attribution-NoDerivs license
Qy	853 FGIVELPLIFTINGSPFAQWNLKLNNSGCFQNNQGRPIFNHSRSRISLGLKK 907	CC	CC
Db	403 PGILGVLCLITVATVAKWLNANEVKIRGN-----PNSASDMSRMSMSKR 451	CC	CC
RESULT 1	04 PHQ2_USTMA PRELIMINARY: PRT: 881 AA.	CC	CC
ID	O4PHQ2_USTMA	Qy	57 SERQOOHQHQETS-----DNKAPLTKRSGSIKKSNLTKDIDTNPMSLSCDDDTI 108
Q4PHQ2_USTMA	PRELIMINARY: PRT: 881 AA.	Db	61 -----SPRASR-----SDQQABPFYBGLNKSPVAVRTAAPPSSAAARY 104
Q4PHQ2_USTMA	04PHQ2_USTMA	Qy	160 -NKSREKSLQENTLPPLIKKKTGRANSNNFENDLV-----SPMTMKNDNSEDTIN 209
04PHQ2_USTMA	04PHQ2_USTMA	Db	105 QHNRSGNTSLDSQHHPDTHQDHSNTHPSALRGSRALSRTPMRSPNTYPPSSRIN 164
04PHQ2_USTMA	04PHQ2_USTMA	Qy	210 TST-----TANHMLKGATLGLGCTGTATAATAGR-----RPSRSSD 252
04PHQ2_USTMA	04PHQ2_USTMA	Db	165 EHTLDDDDAIDODASAHSRGLD-----NNRRSRSHLSSPAYDHDOPTRPSLRBD 220
04PHQ2_USTMA	04PHQ2_USTMA	Qy	233 SEADSHASRSQETEDVCPFM-----VGD-----HRYNG----- 283
04PHQ2_USTMA	04PHQ2_USTMA	Db	221 S-----TVEDDVCPHMIDHTSEGPHETHPEHLDDEVDAYHPRADSSHVPUG 269
04PHQ2_USTMA	04PHQ2_USTMA	Qy	284 ---IDPDEIDFREQEAYLQKOMIA-----KHNLRDEFQNLKUNNTSGA----- 329
04PHQ2_USTMA	04PHQ2_USTMA	Db	270 YPFPTDNLNEPASEDRGVGSSNNMASPPTGTLR-----QREVAQPKGSQAGYGS 325
04PHQ2_USTMA	04PHQ2_USTMA	Qy	330 -----SRHPYTHHSNNNNKNNKNGSDGCGSSEMAALKTPKNIKKTLSRPBPTHE--- 377
04PHQ2_USTMA	04PHQ2_USTMA	Db	336 PDSDDATPKSRP-----RKLSTAGFAGSS-----RK-----ORKLAFLFGESAPP 368
04PHQ2_USTMA	04PHQ2_USTMA	Qy	378 ---NSSSSEIYB-----LKTQOPPYR-----YDQDLSUTSSSSTSNSGCS 416
04PHQ2_USTMA	04PHQ2_USTMA	Db	369 EVAGAAAASSSTMGGGTADAKRSMDVNTPLDNDMLPSYNGNSRMHLSQSQGSLSPVTP 428
04PHQ2_USTMA	04PHQ2_USTMA	Qy	417 QKPGGARTSDG1-----NGGSUPDRPLSPHSSESETHADPSLNSPSGOSVDRPLRNGEE 474
04PHQ2_USTMA	04PHQ2_USTMA	Db	429 GGGMGRARSHTGALSNRGEKTPRAPSMLPSLPHRHSLEPDGTPBELLSGRKE 488
04PHQ2_USTMA	04PHQ2_USTMA	Qy	475 ----- 475
04PHQ2_USTMA	04PHQ2_USTMA	Db	489 AADISPLTESQSNNNARAPTAINTARMSSLQALGRMEGNTAAPTSNSNTPLLAGDVMG 548

Oy	475 -----TWLDCTCPDSEKMLAKAFIHPULATEDIRMQET 510	54 VVKSEKOOHHQEITSNAKPLTRKSGSSSIKKKSNSLTDKDRITNPMSLSGGDTINSGHK 113
Db	549 GUSATRITALPGGRMDPDEANTWMLDVSPTDQENKLILSVFTHPLATEDILMERT 608	63 LLOQVOVDALE-----AGFSIDFQAI-----ADDRSTK-----
Oy	511 REKVLPKSYTVCFHFEADEKEDSELEPUNTVVPHFDGILTTFSPISHPANVRRV 570	114 NRNNYMASSLKDYL-----KONTDNDSTNNHHTHLAIPPIPPTITNNKNSRKSQLENL 171
Db	609 REKIEFRAYHIVCVRSPDQBYSEPTLEPLANNMIVFRETSUHFRTGPQVNRAKI 668	96 ----DNGLGRFLSDPCTNRCGRASFRHOI-----AMNSRDSS-----
Oy	571 RQLRDYDVSAWSWLYALIDEITGQFAPVHGTEYEADAEAVFTARDTDSMMIQRIG 630	172 PPLIKKTGTGNNSNNPENDVSPMTKNTDSDDITNTSTANHMKLG GATTUGVGNG 231
Db	669 KOLKDYINTVNTSWIYALIDDTAFCPLQSTEVYEDSIDENTVLKSAEODMLRIG 728	134 ----VSARSTSPNVEAP----ADPERERANTE-----SHAADPIDEALQRTVSC-GTH 181
Oy	631 ESRKYMULMRLSSKGADVIMCPAKRCOEANSSGYQRYNLUQQQQQAPPNPNTI 690	232 TTATATAAAAGR-RPSRSSIDSEADSHASRSOFTEDVEDCVPMMVGDHIRVNGIDPDETD 290
Db	729 TCRKVMGLRLLAGRADVAGLAKRCNEWS-----760	182 PRPFIFNSAIRFGDQIDS-----PEDCATPFEQLRGIVDTEELIS 228
Oy	691 TSPINSTLNLSLGTSFEGVGSINGPNPNTNTGSPSPQQQQRGITN 750	291 EPTREBEATYLOKOMTAKNLIRDEFQNSKUNTTSGASRHYYHHHSNNNNKANGDGG 350
Db	761 -----760	229 EFV-----ALQOKTKPITTRKHSLSQSCK-----254
Oy	751 KSPPIPDARRADIALYLGDTQDHITMFONLALAYEKIFSRSHSNLTLAQLOVESPNRKK 810	351 GSSMAALKYTPKNLKKTLSPEFTHEHSSSBETVELTKQOPPKYDQLSLTSSTSS 410
Db	761 -----VAPKEDTGLYSDQDLHITQDHLTMQNLNAYEKILSRNSNTLAQISTEMTAHQ 813	255 -----SVFVTDLAPNAQKSAHBAKRSSSEVSESSDDIGLDKDB 295
Oy	811 ITEMFSKITLGMPLVPLNYTGLGRHGNVRVPGEGCTNLGMFFGIVLIFTIGSFPA 870	411 TSGSGSCGVFGGARISGINGSLPDRFLSPHSESEETHADPISLVSPGQSTYDLPR 470
Db	814 INDVLSKLDTAIVTIVPNTVNGTGLGRHGNVPGECVENLRMPAGICSVMVIAVSYTA 873	236 KVFDADYNEK---OLYVENLNEMEKEPREGFPPSSOSVTHAAELDQVPPDPDPLFQ 351
Oy	871 QWALKK 876	471 NGEE-TWLLDCTPOTSEKMMAKARGIPLAEDTRQESTREKBLKSTPYPCHP 528
Db	874 TRYLER 879	352 LGPECGWMLDVLAPTAEVGALSRAFSIMPITLEDITOBANKYBPLQTYVOPRTP 411
Oy	RESULT 15	529 -EADKESDYLEPINVTUVPHDGLUTPHFSPISSHANYRVRVROLQDUDVSADMLYCA 587
ID	QUTU7_JASPOR PRELIMINARY; PRT; 756 AA.	412 YDINKTSERFEPNPYMTIVPROULSPSPTEPHASNYRRIKGURDVSLSMDMCYCA 471
AC	QUTU7_JASPOR	588 LIDETTOOPAVTIGIEVADALDAVPTARDPESMMLORTGESERKNTATMLRSGKA 647
DT	24-JAN-2006, integrated into UniProtKB/TREMBL.	472 MIDDIVDSFGVIVBIEVESEATEDLVIARMDDFESSPUPRIGLAKKYMNSLRIGGKA 531
DE	07-MAR-2006, entry version 3.	648 DVTKMFARCOEANSSGGYVQRQYNIQQQQQAPPNNPNTSPINSTLNLSGFTST 707
GN	Mg2+ and Co2+ transporters.	532 DVTGSPSKRCNB-----543
OC	Eukaryota; Fungi; Ascomycota; Pezizomycetes; Eurotiomycetes;	708 GGGVGVGINFNGPNTPTNTNTGSPPOOOQQGIRTKNSPPIPADPARDIALY 767
NCBI_TaxID	5062; OX	544 -----QVSVT-----PRGDIGV 556
RN	1] NUCLEOTIDE SEQUENCE.	768 LGDIODHITMFQNLLAYEKIFSRSHSNYLAQLOVESFSNNNKITEMFSKTKLIGLVP 827
RC	SRASTRN-IB 40; PubMed=16372010; DOI=10.1038/nature04300;	537 LGDQDHVTTMSTMHLAFKLSLSHTNLYQLNVTNLVGHNVNLKSKVLTATMLVP 616
RA	Machida M., Asai K., Sano M., Tanaka T., Kumagai T., Terai G.,	
RA	Kusumoto K., Arima T., Kitamoto K., Akita O., Kashiki Y., Abe K., Gomi K.,	
RA	Horiuchi H., Kitamoto K., Kobayashi T., Takeuchi M., Denning D.W.,	
RA	Galagan J.E., Nieman R.C., W.C., Ya J., Archer D.B., Bennett J.W.,	
RA	Bhatnagar D., Cleveland T.E., Fedorova N.D., Goch O., Horikawa H.,	
RA	Hosoyama A., Ichinomiya M., Igarashi R., Iwashita K., Juvvadi P.R.,	
RA	Kato M., Kato Y., Kin T., Kokubun A., Maeda H., Maseyama N.,	
RA	Maruyama J., Nagasaki H., Nakajima T., Oda K., Okada K., Paulsen I.,	
RA	Sawano T., Sakamoto K., Takahashi M., Takase K., Terabayashi Y.,	
RA	Wortman J.R., Yamada O., Yamagata Y., Anazawa H., Hata Y., Koide Y.,	
RA	Konozaki T., Koyama Y., Mineo T., Suharnan S., Tanaka A., Isono K.,	
RA	Kuhara S., Ogasawara N., Kikuchi H.,	
RA	*Genome sequencing and analysis of <i>Aspergillus oryzae</i> .*	
CC	Nature 438:1157-1161(2005).	
CC	Copyright by the UniProt Consortium, see http://www.uniprot.org/terms	
CC	Distributed under the Creative Commons Attribution-NoDerivs license	
DR	EMBL: A0007150; BAE55038..1; Genomic_DNA.	
SEQUENCE	756 AA; 85231 MW; 884BC22CD3BA528 CRC64;	
Query Match	20.5%; Score 990; DB: 2; Length 756;	
Best Local Similarity	31.2%; Prod. No. 8.7e-66;	
Matches	Conservative 104; Mismatches 234; Indels 246; Gaps 21;	
Oy	4 SESYQNSTTQIPRSDVLDIRNOI-----TMDCAS-----DSEDBLELKSELSE 53	
Db	7 SEKVKMFLSRFSTPVPE----LDQDRFQDPPRLEATEDMTLSRONTNPMETPQSD 62	

Qy	113 KNR-----NTYNNSSL-----RKDFYLKDNTDNSTANT----HIAPIPPIPPTPI 156	A: Cross-references: UNIPROT:P43553; UNIPARC:UPI00001258DD; EMBL:D50617; NID:g836665;
Db	109 QIRASAITLSAIPSRLAHSNPHORLYVESNI-----HTTPKDYGTYKDYMISSTA 161	PID:d1008822; PID:9836703; MIPS:YFL050C R: Murakami, Y.
Qy	157 TWANKSRKSKOLENLPLPPIKKITGRANNSNNPFNDLVSPIATKMTKTDSDINT---- 210	A: Reference number: S62230
Qy	162 SSGNTS-KLSSSSASAPTKVR-----KSLVSPVLEPHESKEDTHSKLAKPKK 210	A: Molecule type: DNA
Qy	211 ---STTAHMGK-TGATLGTVGTOTATA---TATAAGRRPSRSSIDSADSHASRSQQ 264	A: Cross references: UNIPARC:UPI00001258DD; EMBL:D44603; NID:g871957; PID:d1008660;
Db	211 RYTSYTSASHSIN-PAVLTKTSQKDADDTLERKPYRMNTRAFDSDV-SQASRSQ 268	C: Residues: 1-858 <MUR>
Qy	265 EPEEDYCPMPYGD-HIRUNGIDDEDFKREEREAQYKOKOMIAF-NFLRDEFONLSK 322	A: Cross-references: SGD:50001844; MIPS:YFL050C
Qy	266 DIGFTSSTSCTSS-----DALKYTPR-VSGQGERSESTN 359	A: Map Position: 6L
Db	269 EPEEDYCPMPYGDHTRYNGDPELEYQXOPANAE--KSQPLASQVNEQKYSNTSQ 325	C: Keywords: transmembrane protein
Qy	323 -----NNNTGASASHPYTHHSNNNKNGDGGSSMALXKTPNENLKLRTSLREFTH 376	F: 742-758/Domain: transmembrane #status predicted <TM1>; 777-793/Domain: transmembrane #status predicted <TM2>
Qy	326 DIGFTSSTSCTSS-----DALKYTPR-VSGQGERSESTN 359	Matches 401: Conservative 126; Mismatches 221; Indels 233; Gaps 35;
Qy	377 ENSSSSEELYLKIKTQKPPYKQDQLSLTSSTSSTSGSSGSGYKEGG-ARISDGINGSSL 435	Query Match Score 1666; DB 2; Length 858;
Db	360 ET-----EHHEKEDDEHKIKPSLKHPLISFGKNNKVEBERENIPSDPACYGQTDPQI 414	Best Local Similarity 40.9%; Pred. No. 6-8e-9;
Db	416 PDRSLFHSESEETHADPITSVSPGQSVTRDLPFRNGEETWILDCPCPTDSEKMKLAKP 495	C: Map Position: 6L
Db	415 PNRFFCSEEDYVHASDTSVSEGTQFYLFRGEGPWNIDCSCPDDENRCIATKA 474	Qy 94 RITHNMSLSSGQDDTINSGHKNRY---NMS-SLRKDYLKINTDNTSNMTHLAIPIP 148
Qy	496 GHPLTADIMQETREKVELFKSYTYVCHETBANKESDVKLEPNINXVMPHQGICTP 555	Db 63 R--JHKFSSNSE---NSGVNGGVEKVTNTSSTSGRDF--EGEEAAEVKRYOLSFAL- 114
Db	475 GHPLTADIMQETREKVELFKSYTYVCHETBANKESDVKLEPNINXVMPHQGICTP 534	Qy 149 IPQPPITIANKR-----RKSOLENPLPLIKK-----TIGRNNNSNPEND 191
Qy	556 HFSPISSHANTVRVRQVLDVDSWMLCYALIDETDGAVIDGIVYEADATEDAVF 615	Db 115 -----LSNARSPRLAKSETHQKQHVBESIASPLSKNALEGRHTDLPAGTSSRCLN 168
Db	535 HFGPISHCANYTRVRQVLDVDSWMLCYALIDETDGAVIDGIVYEADATEDAVF 594	Qy 192 LVSPMTNMKMTN-----DSEDTIT-----STTANHMLKGIGATTL 226
Qy	616 TADOTDESSMORIGESERKTMMLRLLSKADTRKPARCDEANGI----- 643	Db 169 EAASSARMTSARKASLVSAIFETSAESENGTIPQAKLKRRTYSVTISTS--SYNPTTL 226
Db	595 MARDDFAMMORIGESERKTMMLRLLSKADTRKPARCDEANGI----- 643	Qy 237 CUGGTTATATAG-----RPRPSRTSIDBANDSHARSQETBEDCPRVG- 276
Qy	676 OOOQQAPPNPNLITSPINTELMLNLSTGSGVGEINGFPPTGNTNTNTTG 735	Db 227 -----LTTAQSQKSDMENDTRPKLQMDSRVPHSPI-SQASRSQETBEDCPRMPLR 280
Db	644 -----GPALTSDQIN-LANLQRQDNA-----SHKNNNSSTTVPRNN 678	Qy 277 DHIRNGGIDFIDEDEFREEREAEYQKQKIANLRLIDEFONLSNNNTSGASRH---P 333
Qy	736 STSPPOOQOOGITNKSPPIFADRADIALYCGDQHMTTQNLAVETKPSRSN 795	Db 281 LHTRNVGTDPELEYA-----QISNAERNISLANHOHSBERT 318
Db	679 APTSQ-----PREDIAYLGQDQHMTQNLAVETKPSRSN 720	Qy 334 YHHHSNNNNKRNKGDGGSSMALKTPKNNLKKTSRFEPHENSSSEBYELKTKQ 393
Qy	796 YLAQLOQESFSNNNNKITEPMSKLTIGTMLVPLNLTGGLKGNVYPERGTTNLGMFPSI 855	Db 319 YNHQDQTCFTNSASTGSS-AALKYTFE-----ISR---TLMRCSNTNEYVSENNES 368
Db	721 YLAQLOQESFSNNNNKITEPMSKLTIGTMLVPLNLTGGLKGNVYPERGTTNLGMFPSI 779	Qy 394 PPKYKDQOLSLSTSSTSGSSGQVKFGGARYSDSINGS-----SLPD 437
Qy	856 VGYLPIITGSSPFTAQWKLKNNSI---EGQNNNHRPFI---HSRSTRSLGRK 907	Db 339 --VREDDRPD-----HDVTFGRNKKEKEGKEDSYSSRAYTQLQTEQIFS 415
Db	780 LAVLILLAVLAVMLASVYMLKIDPPTLNEAASSGAKSVTSSFLPKRKRFRDNRKNTV 839	Db 416 RFSRSPRSDETVHADPSISGQTYFLPGGQPTWMQDSCPDEMRCLAKTFCI 475
Qy	908 HGG--NKSIS1SPNPKY 921	Qy 498 HPLTAEDTRHQETREKVELFKSYXPKCIPPEADKESEDYLPINPVYIVPHDGLTTHP 557
Db	840 RAGPSNKSVASLPSK 855	Db 476 HPLTAEDTRHQETREKVELFKSYXPKCIPPEADKESEDYLPINPVYIVPHDGLTTHP 557
Qy	RESULT 2	Db 538 SPISHANVANRERVQLDYYDVSADWICLALIBDITGAPYTHIEYADIEADAVFTA 617
Qy	55605 Probable membrane protein YFL050C - yeast (Saccharomyces cerevisiae)	Db 536 DPISHCANVERVQLDYYDVSADWICLALIBDITGAPYTHIEYADIEADAVFTA 595
N: Alternate names: hypothetical protein R001	C: Species: Saccharomyces cerevisiae	Qy 618 RDTFSSMQLQIGSERKVKDMLLGRADKXKARAKCQBANSSCYQYRQ---TML 674
R: Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Soeda, E.; Yokoyama, K.; Yamazaki, M.; Tashiro, H.; Eki, T.	C: Date: 02-Sep-1995 #text_change 09-Jul-2004	Db 536 RDMDFAAMLQIGSERKVKDMLLGRADKXKARAKCQBANSSCYQYRQ---TML 655
Sasanuma, M.; Tsuchiya, Y.; Soeda, E.; Yokoyama, K.; Yamazaki, M.; Tashiro, H.; Eki, T.	A: Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces cerevisiae.	Qy 675 QQQQQQAPPNPNLITSPINTELMLNLSTGSGVGEINGFPPTGNTNTNTTG 734
Submitted to the EMBL Data Library May 1995	A: Reference number: S56186	Db 656 QAEQE-----NYNQ-----NNNN-----668
Saccharomyces cerevisiae.	A: Accession: S56205; S62287	Qy 735 GSPSPQQQQHGTNTKNSPPDAPRADIALYCGDQHMTQNLAVETKPSRSN 794
A: Molecule type: DNA	A: Accession: S56205	Db 669 -----NOISLUSN-SYQHTSQPRGDSLALYDQDHLLTMQDNLAYEKPSRSIA 718
A: Residues: 1-858 <MUR>	Qy 795 NYLAQLOQEVSTFSNNNKTEMPSKLTUQTMVPLNLYTGFGRNTVPEGCTNIGMFPG 854	

A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. MUID:8905231	Qy 423 GARISDGINGGSSLIDPRESLPSSESET-IKAPDIPSLVSPGQSYRDL---PRNGEETW 476
A;Accession: S77053	Db 7 GIVSKRKKGKPGSSL---IHVGISTSKTEHIIIDYKDYPERKDVSRSDECAYRDTDNT 63
A;Status: nucleic acid sequence not shown; translation not shown	Qy 477 WLDCPCPDSEM-KMLAKAFGIPHLTAEDTRMQETREKVELFKSYFVCFPHTEAOKES 535
A;Molecule type: DNA	Db 64 WINTGLEDDILYEVKVKYEVHSVHSLMEDILANTQRPKIEEFKFKTFVULKATP---RE 120
A;Residues: 1-380 ;XANS-	Qy 536 DYLEPINYYIVFHFDGLTFLH---FSPSHPANTARRVRQLRDYV-DVSADMCLY 586
A;Cross-references: UNIPROT:Q55968; UNIPARC:UPI00000D344C; EMBL:D64005; GB:AB001339;	Db 121 DKLVVEQVSIICMENLTIFDECGKODGDFV---GNDRDRTNTKSIRTKGVQYLYT 174
NID:91001779; PID:BA10745.; PID:9106532	A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Superfamily: divalent cation transporter, CorA type	Qy 587 ALIDRITDGFAPIHGIEYEADATEDAVFTARDTDFSSMLQRIGESRRKVMTMLLSKK 646
Query Match 4.4%; Score 214; DB 2; Length 380;	Db 175 ALIDSIVDNVYFVILEKDKEFNEESELT-----S 204
Best Local Similarity 19.2%; Pred. No. 1.1e-05; Mismatches 141; Index 140; Gaps 14;	Qy 647 ADVIGMFARCKOEBANNSGGYQROTYNQQQOAAAPPNPNTI---TSPINSTNLMS 702
Matches 89; Conservative 72; Mismatches 141; Index 140; Gaps 14;	Db 205 -----TATNNLFNEVYNLKLQMQ-----IFLWKAWMPREINN--- 236
Qy 478 LDCTCPDSE-----MRLAKAFGIPHLTAEDTRMQETREKVELFKSYFVCFPHTEAOKES 522	Qy 703 LGTNSCGEVGVGGINFGENPQHNTNTNTGSPSPQQQHQGITNKSPPLPDRPRA 762
Db 59 LOCEAVDVESEVSWINIDGLGNHTWEGVPLHVAEDEVNTVQPKVVEYENHLI 118	Db 237 --TQRGEVET-----S 248
Qy 523 VCPHPEADKSESDYLEDPEINYVTVYVENDGILTEFHSP-ISHPAVNRBRVQLRDYV-DVS 580	Qy 763 DIALYLGDOIQDHITMFONLLAYERKIFSRSHSYLAQLOVESFSNNKITEMPSKTLIG 822
Db 119 FISRMVATLQSSQPTISQFISLQRKH-YLTIDIEPKYDLSVBSRERLTKGAIQRN 177	Db 249 DISIYKQVDTHTQVQIDTMFAGDGTDL-----SASNKANEMLKPTTFRS 301
Qy 581 ADMLCYALIDESTDGPAPIHGIEYEADATEDAVFTARDTDFSSMLQRIGESRRKVMTLM 640	Qy 823 TMLPMLNVTGLFEGMAYTRVPOEGTNLGMFPGCIVGLULIGSIFPAQMILKK 876
Db 178 ADMLCYALIDESTDGPAPIHGIEYEADATEDAVFTARDTDFSSMLQRIGESRRKVMTLM 640	Db 302 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
Qy 641 RLLSGKADYKMFARCKOEBANSSSGTYQROYNQQQOAAAPPNPNTIPISTNLMS 700	Db 351 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
Db 235 RAIPWPODAI-----244	Db 352 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
Qy 701 NSLGIGSTGCGEVGGINFGPNPTGRNTNTNTGSPSPQQQHQGITNKSPPLDARP 760	Db 353 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
Db 245 NSL-----257	Db 354 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
Qy 761 RADIALYLGDOIQDHITMFONLLAYERKIFSRSHSYLAQLOVESFSNNKITEMPSKTL 820	Db 355 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
Db 258 --EVYTFURDCTDHTQIQLDTIETYDILTVYSSY---SHRMNEIMKLTIV 308	Db 356 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
Qy 821 IGMALVPLNLYTGLFGMAMVRYPGSGTIN--LGMFFG--IVGVLFITLGSFTAQN--- 872	Db 357 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
Db 309 ISSIPPLTPIYGTYGAFNPKRSRPNMPBLANKWYGYTPWMMLVGMMLYFWRKG 367	Db 358 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
Qy 873 WLRKUNNSTEGGNNR 889	Db 359 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
Db 368 WPRNLND---VEKRN 380	Db 360 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
RESULTS 10	Db 361 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
E97128 magnesium and cobalt transport protein CAC1852 [Imported] - Clostridium acetobutylicum	Db 362 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
C;Species: Clostridium acetobutylicum	Db 363 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 31-Dec-2004	Db 364 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
C;Accession: E97128	Db 365 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
F.Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, H.M.; Dubois, J.; Qiu, D.; Wirth, J.; Wolf, Y.I.; Titusov, R.L.; Sabathe, P.; Douette-Stamm, L.; Souaille, P.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.; Bacteriol. 183, 4823-4838, 2001	Qy 366 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
A;Reference number: A96900; MUID:21359325; PMID:21359325	Db 367 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
A;Accession: E97128	Db 368 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
A;Status: preliminary	Db 369 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
A;Molecule type: DNA	Db 370 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
A;Residues: 1-354 ;XURP-	Db 371 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
A;Cross-references: UNIPROT:Q97104; UNIPARC:UPI00000CA31D; GB:AE001437; PID:G1524829; GSDB:GN00168	Db 372 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
A;Experimental source: Clostridium acetobutylicum	Db 373 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
C;Genetics:	Db 374 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
A;Gene: CAC1852	Db 375 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
C;Superfamily: divalent cation transporter, CorA type	Db 376 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
Query Match 4.4%; Score 212.5; DB 2; Length 354;	Db 377 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
Best Local Similarity 18.8%; Pred. No. 1.3e-05; Mismatches 159; Index 149; Gaps 15;	Db 378 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
Matches 89; Conservative 77; Mismatches 159; Index 149; Gaps 15;	Db 379 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
RESULTS 11	Db 380 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
AG2319 hypothetical protein al1410 [Imported] - Nostoc sp. (strain PCC 7120)	Db 381 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
C;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120	Db 382 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
C;Accession: AG2319	Db 383 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
R.Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchi, M.; Ishikawa, T.; Kishida, Y.; Kohara, M.; Matsunoto, M.; Murakami, A.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yesuda, M.; Tabata, S.	Db 384 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
DNA Res. 8, 205-213, 2001	Db 385 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120	Db 386 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
A;Reference number: A81807; MUID:21595285; PMID:11759840	Db 387 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
A;Accession: AG2319	Db 388 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
A;Status: preliminary	Db 389 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
A;Molecule type: DNA	Db 390 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
A;Cross-references: UNIPROT:Q8YPT1; UNIPARC:UPI00000CEA09; GB:BA000019; PID:BA175809; PID:17133245; GSDB:GN00179	Db 391 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
C;Genetics:	Db 392 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
A;Gene: A81807	Db 393 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
C;Superfamily: divalent cation transporter, CorA type	Db 394 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
Query Match 4.3%; Score 207; DB 2; Length 380;	Db 395 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
Best Local Similarity 20.0%; Pred. No. 3e-05; Mismatches 126; Index 144; Gaps 18;	Db 396 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
Matches 85; Conservative 71; Mismatches 126; Index 144; Gaps 18;	Db 397 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
C;Superfamily: divalent cation transporter, CorA type	Db 398 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
Qy 479 WLDCPCPDSE-MCMMLAKAFGIPHLTAEDTRMQETREKVELFKSYFVCFPHTEAOKES 534	Db 399 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
C;Accession: AG2319	Db 400 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
R.Yamada, T.; Nakamura, Y.; Watanabe, A.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchi, M.; Ishikawa, T.; Kishida, Y.; Kohara, M.; Matsunoto, M.; Murakami, A.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yesuda, M.; Tabata, S.	Db 401 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
DNA Res. 8, 205-213, 2001	Db 402 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120	Db 403 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
A;Reference number: A81807	Db 404 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
A;Accession: AG2319	Db 405 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
A;Status: preliminary	Db 406 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
A;Molecule type: DNA	Db 407 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
A;Cross-references: UNIPROT:Q8YPT1; UNIPARC:UPI00000CEA09; GB:BA000019; PID:BA175809; PID:17133245; GSDB:GN00179	Db 408 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
C;Genetics:	Db 409 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
A;Gene: A81807	Db 410 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
C;Superfamily: divalent cation transporter, CorA type	Db 411 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
Query Match 4.4%; Score 212.5; DB 2; Length 354;	Db 412 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
Best Local Similarity 18.8%; Pred. No. 1.3e-05; Mismatches 159; Index 149; Gaps 15;	Db 413 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351
Matches 89; Conservative 77; Mismatches 159; Index 149; Gaps 15;	Db 414 TTFIBLSPLVGVGMNFDMPE---LKFRYGAIVLWIMISLAFMAYFRK 351

Oy	552	ILTPHFSPLISHIPANVRVQLRDYDVDSADWLYALIDBITHGPIAVHGEYEADAE	611
Db	541		
Oy	612	DAVPTARDTDFSSMLQIGESRKVMTMRLSCKAD-----	654
Db	555	-----NLSEEQKINE--KTEBENLTSNSNSNSNSNSNSNSSTISCNPLSER	602
Oy	657	CQEANSGGTQROYNQQQQAPPNNPPI-TSPINSTLNLSGT--SGGGVG	712
Db	603	SKNPSSSTSNTSNTSLRPSQELTEIRPLRAELYSSINTSPRSSASLSCGSSGC	662
Oy	713	VEGIFGNPNPGNNTNT-----NTGSPPOQQOOGITNKSFPIPDARPRAD	763
Db	663	GCGNNNGCPEKSSNSNSPIQFENENESIDSTEKKNEQFESLIPOLI-----RE	713
Oy	764	IALLYGDTQDHITMFQNLAYEKIFSPRSNSNLAQOV-----ESPNRNRK	810
Db	714	NQLYTKPTE-----FREIKLELENSNSNSNWQIEYKSTOLVLKOPKDQDSDNTEK	767
Oy	811	ITEMFSKITLIGTMUPLNLYGLFGHNVRVPGEGSTNLGWFFGTVGVLIPIIGSFIFA	870
Db	768	RQLEN-----GSNV-----SGSN-----	781
Oy	871	QWIKRLANNSEGQNGNRPINISRRSTRSLRKHGONKSIIFPKRY	921
Db	782	-----NGSGGRNNHHHCINNS-----NGSNSEVI--PSKY	811

Search completed: June 13, 2006, 14:09:04
 Job time : 52 secs